DOCUMENT RESUME

ED 186 492 TM 800 245

AUTHOR Millsap, Mary Ann; And Others
TITLE Women's Studies Evaluation Handbook.
INSTITUTION National Inst. of Education (DHEW), Washington, D.C.
PUB DATE Dec 79
NOTE 66p.

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Data Analysis; *Evaluation Methods; Higher Education; Measurement Techniques; Objectives; *Program Evaluation; Quasiexperimental Design; Research Reports; *Womens Studies

ABSTRACT

The purpose of this handbook is to encourage and assist women's studies directors and faculty to perform or obtain evaluations of their programs in order to assist in their decision making needs, and to know how, when, and to what ends program evaluation can be used. Evaluation is defined and described. Five major approaches are listed: professional judgment, measurement, decision-theoretic, goal-free, and the fit between goals and reality. Goals and objectives are defined and possible dilemmas between long- and short-range goals discussed. A quasi-experimental design is described and exemplified for a program with seven objectives. Survey instruments, interviews, and observation techniques are described and compared. The need for planning data analysis in advance of data collection is presented. The final evaluation report is discussed with reference to its audiences (the client and other interested persons) and its format. An example of an executive summary report is appended, together with abstracts of 13 evaluations of women's studies programs, and a list of resource materials. (CTM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED186492

WOMEN'S STUDIES

EVALUATION HANDBOOK

TM800 245

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

The National Institute of Education

WOMEN'S STUDIES
EVALUATION HANDBOOK

December 1979

Mary Ann Millsap
Naida Tushnet Bagenstos
Margaret Talburtt

U.S. Department of Health, Education, and Welfare
Patricia Roberts Harris, Secretary
Mary F. Berry, Assistant Secretary for Education

National Institute of Education
Michael Timpane, Acting Director
Program on Teaching and Learning
Lois-ellin Datta, Acting Associate Director
Testing, Assessment, and Evaluation
Jeffry Schiller, Assistant Director

The contents of this report do not necessarily reflect the views of the National Institute of Education or of any other agency of the U.S. Government.

NOTE TO READERS

The spirit of this handbook is one of enlightened self-interest. Its purposes are to help women's studies directors and faculty realize that no single approach to evaluation is sacrosanct and to show how evaluation can serve their decisionmaking needs. In this sense, the goal is to educate the consumer to know how, when, and to what ends program evaluation can be used.

The handbook provides information on:

- Choosing among types of evaluation and using evaluation;
- Turning program goals into objectives;
- Selecting appropriate research designs;
- Deciding on instruments and items;
- Analyzing data;
- Reporting results;
- Using other materials (summaries of recent studies and resource materials).

This handbook grew out of an evaluation workshop given at the National Women's Studies Association (NWSA) First National Conference in Lawrence, Kansas, in June 1979. The three leaders of the workshop and the authors of this handbook are: Mary Ann Millsap, NIE Program on Teaching and Learning; Naida Tushnet Bagenstos, NIE Program on Dissemination and Improvement of Practice; and Margaret Talburtt, Formative Evaluation Research Associates (Ann Arbor, Michigan).

Early in 1979, Dr. Elaine Reuben, coordinator of NWSA, had asked us to conduct the workshop because of increasing demands for evaluation services, growing concerns about what evaluation is, and questions about how it could best be used. To many, evaluation was seen as a baffling, threatening, yet necessary exercise. Drawing on our own experiences in evaluating innovative educational programs and products, we hope to remove some of the mystique of evaluation and provide directors with a useful tool for assessing their programs. The handbook reflects our commitment to the improvement of education for student development and our belief in the positive uses of evaluation toward that end.

Evaluation is not a mystical or unknowable force but rather a pragmatic tool for program improvement.

This handbook also reflects the philosophy of the NIE Testing, Assessment, and Evaluation Program, which seeks to expand the technical roles of evaluators to include more interaction with stakeholders—with those persons having an investment in the program and its objectives. Through a variety of NIE projects, especially in urban education, evaluators are working directly with school, community, and parent groups in the design, conduct, and use of program evaluations. This handbook, with its emphasis on evaluation and its uses for women's studies directors, is another effort to translate that philosophy into practice.

We are particularly indebted to Elaine Reuben for making the workshop and subsequent handbook possible. To the women's studies directors from colleges and universities all over the country who participated in the workshop, we are also indebted. They served both as a field test for the materials and a support group for revisions. Jane Siegel of the American Institutes for Research provided us with both materials and suggestions for conducting the workshop and preparing the handbook. Had she not been giving a workshop on evaluation for the Women's Educational Equity Grants Program, she would have been the fourth workshop leader at the NWSA conference. To the many pioneers in women's studies and in evaluation who reviewed earlier drafts of the handbook, we are grateful for their thoughtful comments on content, sample items, and additional references. These people include Florence Howe and Paul Lauter of SUNY's College at Old Westbury, Holly Knox of the Project on Equal Education Rights, Meredith Larson of SRI International, and C. B. Crump, formerly of NIE's Social Process/Women's Research Program and now at the University of California at Berkeley.

The handbook is a living document, designed to be revised as we learn more about women's studies and evaluation. We encourage readers to contribute to the content and to comment on the usefulness of the handbook. Please send your suggestions to Mary Ann Millsap, Program on Teaching and Learning, Testing, Assessment and Evaluation Program, National Institute of Education, 1200 19th Street, N.W., Washington, D.C. 20208.

Mary Ann Millsap, Senior Associate
National Institute of Education

CONTENTS

	Note to Readers	iii
1.	Evaluation and Its Uses	1
2.	Goals and Objectives	13
3.	Research Designs	17
4.	Instruments and Items	27
5.	Data Analysis	35
6.	Reporting Results	39
	Appendix A: Example of Executive Summary	41
	Appendix B: Abstracts of Studies	47
	Appendix C: Resource Materials	54

1

EVALUATION AND ITS USES

Introduction

Evaluation is essentially a process of assigning value, of trying to determine the worth of some thing or some activity. Each of us conducts our own evaluations every day. To say that I did better or worse than I expected on an exam or in giving a workshop is an evaluative statement. In assigning grades to students, one looks at whether the students learned as much as or more than some expected average. One makes judgments using information about the meeting of a criterion (or of criteria) and assigns grades.

Program evaluation differs somewhat from such everyday evaluations. In judging a program, one tries to be explicit about what its objectives are, what measures are going to be used to see whether the objectives have been met, and what factors seem to be responsible for creating the effects.

Over the past few years, a mystique has developed about what evaluation is. We need to understand a little more about this mystique before we examine how evaluation can be of use to women's studies programs. Twenty years ago, evaluations consisted of the review of a program or its personnel by a panel of experts, persons whose informed judgment was employed to examine whether the proper range of programs was offered, whether the work was organized in an efficient manner, and whether attitudes about a program were positive. The use of expert consultants in site reviews fulfills a needed function, and such consultants have been used by a number of women's studies programs. These expert consultants are aware of what is going on in different programs and can provide a broader perspective than program personnel typically have. Their reports can offer recommendations for improvement and, in some cases, can provide a "seal of approval" for a program.

Within the past 15 years, evaluations have taken on a wider focus and incorporated a number of new mechanisms. The newness of the field, in fact, is partially responsible for the evaluation mystique. When many of us were students in college, there were no courses in evaluation. Many of the people currently in the field have no formal university

training in evaluation but have academic training in the social sciences (such as in survey research in sociology or in experimental design in psychology). Even now there is often no one university department where one can go to find out what evaluation is. Consequently, many questions about the terms used in evaluation (such as tests of significance, analysis of variance or covariance, true control groups, and quasi-experimental design) remain unanswered. A language has developed for which no lexicon is readily available. In addition, as courses and programs in evaluation have grown, people have developed models (such as formative evaluation, summative evaluation, the decision-theoretic approach to evaluation, goal-free evaluation, discrepancy evaluation), and these models are not always clearly described.

In addition to confusion about models and terms of evaluation, other difficulties arise when one explores the ways in which evaluation is used or can be used. Evaluations designed to meet the specific information needs of sponsors may be quite narrowly focused. Evaluations often are used, however, by individuals or interest groups concerned with objectives that the evaluation may not have addressed. In some cases, the absence of information is equated with negative information to the detriment of programs. For example, while evaluations are often assumed to be concerned only with a program's effects, evaluations also can look at the process of a program's development in order to provide diagnostic information to program managers. Such process evaluations need not address the impact of a program at all. Yet some audiences, upon discovering that program effects were not addressed, may conclude that the program was unsuccessful. Also, an evaluation focused entirely on outcomes, such as measures of student achievement, may not address adequately the quality of a program if no steps were taken to determine whether the results were directly attributable to the program and not to other factors.

It is not surprising, then, that program personnel often perceive evaluation as a threat; they assume that if evaluations do not show programs doing all the things they want to do, the programs will be punished by having funds withdrawn. This concern is not without foundation, and evaluations have been used to limit or reduce funding. While some programs have legitimately lost funding when they were not successful in attaining many objectives, evaluations often have been too narrowly defined to fully explore what programs are doing. As a result, these evaluations have presented an incomplete picture of programs' accomplishments.

The potential or perceived threat of evaluation may have a legitimate political base. Within much of elementary and secondary education, for example, innovative programs but not traditional curriculums are evaluated. This also has been the case with women's studies programs in colleges and universities. The traditional liberal arts curriculums seldom are required to state explicitly their objectives for students, yet

women's studies programs not only are expected to do so but also are judged by whether they have met those objectives. An additional burden often placed upon innovative programs is that they are seen as the panacea for the ills of higher education. Hence, unrealistic expectations are formed, sometimes by program personnel but more often by the institution. As a result, innovative or interdisciplinary programs frequently must contend with unrealistic objectives, competing with each other for funds and support, while traditional offerings continue without examination or question.

In addition, unrealistic expectations can contribute to a premature evaluation of the program or to an evaluation based on insufficient evidence. While we have less information about women's studies programs, some innovative elementary and secondary education ideas (e.g., educational vouchers) lost support after limited tryouts. The conclusions drawn from the few cases were generalized to apply to all such efforts. The result was that the ideas were not tried out in a variety of settings to determine whether they could prove effective in a different environment.

While one can argue with the decisionmaking process and about what is going to be evaluated, it seems likely in a period of retrenchment in higher education that women's studies programs and other interdisciplinary programs are going to be evaluated. It is seemingly no longer a question of *whether* women's studies are going to be evaluated, but *how* they are going to be evaluated. Consequently, women's studies directors and faculty need to become both knowledgeable consumers and active producers of evaluation information.

Types of Evaluation

There are five major approaches to evaluation, each with a very different purpose and methodology.[1] Each is listed with a brief description below. We should keep in mind that while these approaches differ, it is possible and often desirable to combine them in evaluating programs.

Professional judgment. The major purpose of this mode of evaluation is to certify the acceptability of a program as seen by an external group of colleagues using pre-established reference points. Examples of this type of evaluation include accreditation visits and other professional certification of educational programs. The methodology frequently involves curricular reviews, interviews, and a vast collection of supporting materials. The result is professional acceptance, rejection, or—sometimes—a provisional or temporary rating. While often conducted by external groups, this kind of evaluation can also be done through in-house review,

[1] Don Gardner, "Five Evaluation Frameworks," *The Journal of Higher Education*, September/October 1977, Vol. XLVIII, No. 5, pp. 571–593.

involving, for example, people from other interdisciplinary programs. This probably remains the most typical and most well-known form of evaluation in higher education.

Measurement. One example of this evaluative approach is grading. Students' work is measured and compared, and a final evaluative assessment is made. Other examples are personality tests, Graduate Record Exams, the College Boards, and such institutional measures as the Institutional Goals Inventory. In all of these cases, the tests are preestablished, applied to all in an equal manner, and analyzed to look for spread or for differences among individuals. The object of the study is often the student rather than the program. In fact, the central controversy about evaluation as measurement or testing is that the tests often are based on what other students learned in the past rather than on what was taught within the particular college or university.

Decision-theoretic approach. This form of evaluation attempts to involve all decisionmakers in the design of the evaluation. In women's studies, Marcia Guttentag pioneered the use of the decision-theoretic approach.[2] All decisionmakers (e.g., students, faculty, administrators, funding agencies) are asked to list what they think are the objectives of the program, to rank the objectives numerically in order of importance (so one can then say, for example, that some objectives are 2 or 10 times as important as others), and then to estimate the chances of the objectives being accomplished (these are called prior probabilities, and a score of 0.5 would be a 50-50 chance while 0.9 would mean virtual certainty). Data then are collected on each objective and analyzed through the use of Bayesian statistics. The most useful part of the decision-theoretic approach is the deliberate involvement of all decisionmakers prior to the conduct of the evaluation. Several other models, such as Provus' discrepancy evaluation, also rely heavily on the early involvement of potential decisionmakers.[3] While time consuming and difficult, this early involvement can lead to the development and acceptance of more realistic objectives for a program, to a more sensitive evaluation design, and to greater use of the evaluation data in decisionmaking.

---

[2] Marcia Guttentag, ed., *Evaluation Studies Review Annual*, Vol. 2 (Beverly Hills: Sage Publications, 1977). Marcia Guttentag, Lorelei R. Brush, Alice Ross Gold, Marnie W. Mueller, Sheila Tobias, and Marni Goldstein White, "Evaluating Women's Studies: A Decision-Theoretic Approach," *Signs*, Vol. 3, No. 4 (Chicago: University of Chicago Press, Summer 1978), pp. 884–890.

[3] Malcolm Provus, "Evaluation of Ongoing Programs in the Public School System," *The 68th Yearbook of the National Society for the Study of Education*, Part II (Chicago: The University of Chicago Press, 1969), pp. 242-283.

Goal-free evaluation. The important aspect of this form of evaluation is that it pays as much attention to the unintended consequences of a program as to intended outcomes, and every program has both. The intended results are those embodied by the goals and objectives. But unintended consequences can be equally important. For example, the Career Facilitation workshops of the National Science Foundation's Women in Science Program are designed to provide information about potential careers in science to young women. One result found by evaluators, however, was that the development of the workshops led to the creation of strong and supportive networks among women who were *already* scientists. An evaluation that searches for unintended consequences is called goal free because it is not *bound* by the program's goals. Since many positive aspects of programs—and some of the more troublesome, negative aspects—are unintended, evaluation designs might, insofar as possible, look for all effects, goal related or not. The methodology of goal-free evaluation is almost anthropological. One would examine the artifacts of a program, talk with its living survivors, and review the records, all the while checking to see what happened and why rather than whether objectives were met. Major proponents of this approach are Michael Scriven and Malcolm Parlett.[4]

Assessment of the fit between goals and reality. This form of evaluation is probably the approach that most innovative programs use. The purpose of the evaluation is to understand whether the stated objectives of a program were achieved. The methodology can range from a classic experimental design to informal interviews.

With the exception of goal-free evaluation, all of the models described above also can incorporate an examination of what was accomplished in light of what was expected. However, because the assessment-of-fit model could in fact be an overarching one, much of the rest of the handbook addresses its application to women's studies programs.

Uses of Evaluation

The conduct of evaluations, particularly the last three discussed above, can be useful to people directing women's studies programs. They can provide information on the extent to which programs are attaining their objectives, on where strengths and weaknesses are, and on what areas require improvement. Moreover, evaluations can buttress other kinds of information that women's studies programs probably ought to be col-

[4] Malcolm Parlett and Garry Dearden, eds., *Introduction to Illuminative Evaluation* (Pacific Soundings Press, 1977). Michael Scriven, "The Methodology of Evaluation," in *Perspectives of Curriculum Evaluation*, R. W. Tyler, ed. (Chicago: Rand McNally, Inc., 1967), pp. 39–83.

lecting as well. For example, while evaluations often are concerned with the impact of programs on students or faculty, such additional questions as the following should be posed.

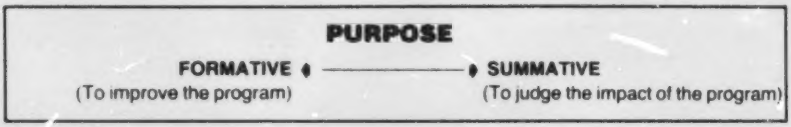
- Has there been an increased demand for women's studies programs (e.g., has the unduplicated count of people enrolled in women's studies increased, are more students anticipating taking women's studies programs)?
- Have the programs increased the amount of external money coming into colleges and universities?
- Have more students been attracted to the college or university in part because of the women's studies programs?
- Have students enrolled in women's studies programs had higher retention rates in colleges and universities than other students?

These are all forms of unanticipated consequences, discussed under goal-free evaluation.

For women's studies programs to compete successfully for the allocation of resources and staff, such information could be extraordinarily useful. In fact, women's studies directors should be encouraged to view evaluation in an active rather than a passive light. Programs frequently can be criticized for not having sufficient evaluation data, but if programs have solid, positive results, the data can be used not only in discussions with the institution but also in securing funds from other public and private sources.

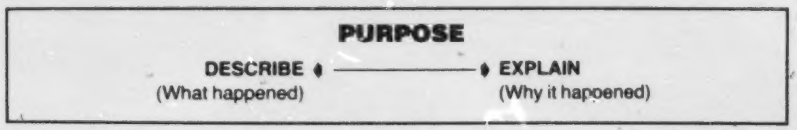
Critical Characteristics of an Evaluation

The time, money, and effort put into a program evaluation will be enhanced if the complete range of alternatives is considered before the research begins. By linking decisionmaking needs to research questions at the start, greater impact of the evaluation is ensured. The following discussion presents the characteristics that we feel program directors should consider in developing evaluations. These points are particularly relevant in a goals/reality evaluation, but could be useful in the design of other approaches to evaluation as well. Each of these characteristics is expressed as a continuum, but most evaluations choose a balance somewhere between the two extremes. The availability of time, money, and staff is a necessary consideration in designing an evaluation. The more formative, comprehensive, and thorough an evaluation is, the higher its financial, temporal, and staffing costs.

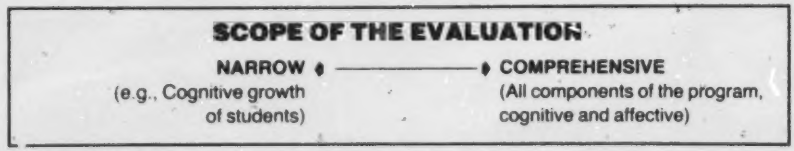


A formative evaluation is meant to collect information that improves the program. Collection usually starts soon after the program begins and is ongoing. It is the type of evaluation a program director often wants, and is sometimes referred to either as a process or management evaluation.

A summative evaluation judges the program's merit. Was it worth the money? Is it qualitatively acceptable? Should it be continued or expanded? Certainly these considerations are important to the program directors as well as to those in decisionmaking positions above them. Summative evaluations usually come toward the end of a program and generally affect the fate of a program. For this reason, summative evaluations often are feared and are confused with other types of evaluation. Most programs want to do some of each; the real issue is where to stand between the two.



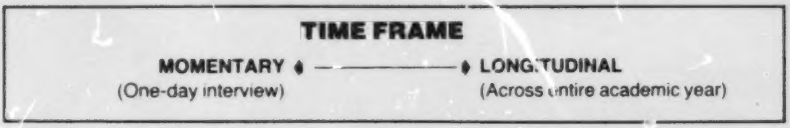
Similarly, one needs to decide what amount of the evaluative work will *describe* impact (e.g., 75 percent of all students answered at least 80 percent of the items correctly) and how much of the study can *explain* impact (e.g., students answered the questions correctly for these three reasons). To determine the appropriate position to take on these two purpose continua, the program director should consider such points as why and for whom the evaluation is being done.



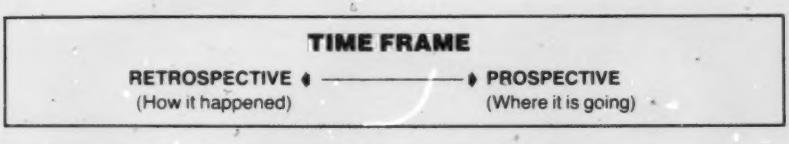
Does the program seek to have an impact on only one element, such as student awareness of women authors in the United Kingdom? Or does the program hope to affect multiple cognitive and affective variables? Are

students in class x to be compared to students in class y? To students in other schools? The more goals a program has, the more comprehensive its evaluation should be.

One may choose to evaluate the objectives one by one or in a closely integrated design. The greater the scope of evaluation, the higher its costs, but the more likely it is to capture the richness of a program and its effects. Conversely, a small program with limited resources for evaluation would be well advised to focus on a few objectives and keep its evaluation appropriately refined.



An evaluation can take place at the end of the semester, at multiple points during the semester, or even long after the semester is over. By monitoring a program over time, one is in a much stronger position to understand its evolution and impact. The monitoring can be simple or complex, depending on the scope, purpose, and resources of the evaluation. The program director should decide before a program starts if the longitudinal information is desirable. If data are not obtained before or as a program starts, one cannot go back to retrieve them after the program is underway, since by that time one no longer can distinguish between what was known before and what was learned during the program.



A retrospective evaluation seeks to explain with hindsight why a project evolved as it did. This type of information often is required by funding agents who want to know why their money was spent as it was. On the other hand, a prospective evaluation may be of greater interest to a program director who wants to determine options for the future. An evaluation can study both of these issues if desired. Again, the real decision is one of emphasis.

DESIGN AND IMPLEMENTATION

CONDUCTED INTERNALLY (Within department) ◄——————► **CONDUCTED BY EXTERNAL PARTY** (Outside researchers)

Before an evaluation begins, one must know who is doing it. Many options exist. Program directors may choose to do all evaluation themselves. This poses some problems in terms of credibility and sheer stamina. Some administrators, citing potential vested interests, may not believe the findings of an internal evaluation, especially one of program effects. In contrast, it is possible to hire someone from an outside firm to conduct all evaluation activities. Between these two points are many options. A program director may wish to hire an outside evaluator to design a study and instruments, but carry it out internally. A program director could hire an outside expert in the field to review parts of the program as appropriate (e.g., a noted English professor to comment on the curriculum of "Women Authors 401"). Or, some universities may have in-house researchers available.

Selection of an evaluator is complicated by issues of political, economic, and professional competency. Is it better to have an evaluator who knows a program well or not at all? How much money is available for outside resources? Who is credible to those who commissioned or mandated the evaluation? Will the evaluator(s) be evaluated? Regardless of who conducts an evaluation, a program director would do well to review the purposes, scope, time frame, design and implementation, methodology, and feedback with all candidates for the job. Do not accept an evaluator who refuses to discuss the parameters of the evaluation in plain English.

METHODOLOGY

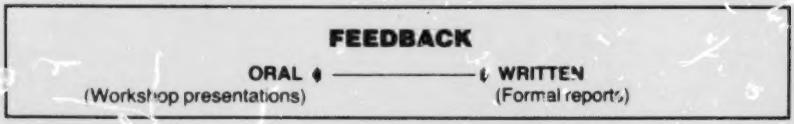
INSTRUMENTED (Quantified information) ◄——————► **OPEN-ENDED** (Qualitative information)

There are many different and legitimate ways to collect data.[5] The purposes of the evaluation, the acceptability of the methodologies to those involved, the necessary costs and time restraints all are to be taken into

---

[5] Scarvia B. Anderson, Samuel Ball, Richard T. Murphy and Associates, *Encyclopedia of Educational Evaluation* (San Francisco: Jossey-Bass Publishers, 1973 & 1975). Sara Steele, *Contemporary Approaches to Program Evaluation*, ERIC Clearinghouse on Adult Education (Washington, D.C.: Capitol Publications, 1973). Elmer L. Struening and Marcia Guttentag, eds., *Handbook of Evaluation Research*, Vols. I & II (Beverly Hills: Sage Publications, 1975). Daniel L. Stufflebeam, et al., *Educational Evaluation and Decision Making* (Itasca, Illinois: F. E. Peacock Publishers, Inc., 1971).

account in deciding this point. Are packaged tests to be used? Would interviews be preferable? Will surveys be returned if they are sent out? Should they be open-ended or confined to quick checklists? Can a program director review or shape questionnaires or interviews? The importance of even a small pre-test of pilot data collection cannot be stressed enough to determine whether the instruments really probe for what the evaluators want to know. Frequently the best methodology is a combination of approaches, permitting the inclusion of all types of data from all types of sources.



This is often the most ignored aspect of evaluation. The point is to analyze and report the findings so that they can be used rather than forgotten. For some audiences, a 200-page report is a must and will be thoroughly utilized. But, for others, a two-page executive summary, highlighting the major findings, will have far more impact. Perhaps a workshop or 1-hour oral presentation would gather more attention. A video tape may document the results better than a report. In choosing the mode(s) of feedback, consider the learning styles of the audience and what they would view as credible *and* interesting documentation of a program's impact.

Summary

This chapter is intended to dispel some of the myths surrounding program evaluation and to replace them with a basic knowledge of evaluation. From the five general approaches to evaluation we moved into an exploration of the parameters that need to be considered in the design of an evaluation. As a summary note, the following questions translate many of the principles of evaluation into some very practical questions. By reviewing these points early in a program's life, directors will be able to formulate an evaluation plan that best serves their needs as well as to prepare for the summative judgments others will make.

Practical Steps for Conducting Evaluations

A. Identify evaluation needs and purposes.

Who should be involved?
Who should decide the scope of the evaluation?
What process is best suited to making decisions?

B. Create research questions.

What must be studied?
What additional information would be desirable?

C. Create evaluation design.

What kinds of evidence are necessary and credible?
What financial and personnel resources are available?

D. Create evaluation budget.

How much time will be involved for various personnel?
What salary and direct and indirect costs are involved?

E. Prepare evaluation working plan.

What are the specific evaluation tasks?
Who should do them?
When should they be accomplished?

F. Collect evaluation information.

What are the sources of information?
What ethical questions are involved?

G. Prepare information for analysis.

What computer-related resources are necessary and available?

H. Conduct information analysis.

What analytical methods are appropriate?
What interpretations can be made, given the strengths and weaknesses of the study?

1. Report the information.

   Who are the audiences?
   What are their decisionmaking needs?
   What are their learning styles?

The remaining chapters address in more detail the practical steps to consider in conducting an evaluation. These include translating goals into objectives and research questions, creating research designs, selecting instruments and items, analyzing information, and reporting the results.

2

GOALS AND OBJECTIVES

Since evaluation involves judgment against some standard, it is frequently useful to judge a program in light of its own goals and objectives. As discussed in the first chapter, measuring a program in terms of its objectives is only one approach to evaluation. It is, however, an extremely common one and one that is generally useful to decisionmakers. In this chapter, the process of developing statements of goals and objectives is described; if a program is to be judged in terms of whether the goals and objectives are accomplished, they must be clearly stated. The process of making the goals and objectives explicit can do more than guide the evaluation. It can help faculty and students focus on what *they* want to accomplish and guide program development so that each component (e.g., course, special event) can contribute to the desired end.

Definitions

Goals are generally long-term, fairly global aims. A goal statement concerning women's studies from the University of Massachusetts, for example, is:

> . . . the development of frameworks and methodologies which integrate women's experiences and scholarly disciplines, and which ultimately will provide the incentives and expertise necessary to effect change in the larger community.[1]

Objectives are usually more specific and attainable. Some evaluations also describe them as measurables. At the very least, there must be a direct way to gather evidence about their achievement. Often the

[1] "Women's Studies Goal Statement," University of Massachusetts, Amherst, Massachusetts, 1976 (referenced in *Undergraduate Women's Studies Program* by Carolyn Rhodes, Old Dominion University Research Foundation, November 1978, p. 58).

method for evaluation is included in the objective. Two types of objectives—*impact* and *process*—are important. As the names imply, the former looks at results while the latter looks at the ways used to achieve them. An example of an impact objective is:

> Students in pre-service teacher education courses shall apply an analytical scheme for assessing sexism in textbooks. (Method of evaluation: Given a textbook, 75% of the students will select and use correctly an appropriate analytical scheme to assess sexism.)

And a related process objective could be:

> to develop and teach a section of the Foundations of Education course that presents methods of analyzing sexism in textbooks. (Method of evaluation: Analysis of course syllabus.)

An evaluation of the program logically might show that the process objective was achieved (the section of the course was developed and taught) but the impact objective was not (students couldn't analyze the textbooks for sexism). In that case, the staff and students would need to look further to discover why the course was not successful and, as a result, would probably want to make program adjustments.

Developing Statements of Objectives from Statements of Goals

In this section, a general goal statement is presented along with a few examples of objectives that are logically related to it.

> As part of the Feminist Movement, Women's Studies courses should function: (1) to help women examine alternative ways of looking at their roles in society and the assumptions of our culture, and (2) to discover and provide new information on the history, culture and accomplishments of women. . . . In addition to its interdisciplinary nature, what makes Women's Studies different from most other academic subjects is that it attempts to foster affective as well as cognitive learning.[2]

[2] Jane Howard, *Liberating Our Children, Ourselves*, 1975 (referenced in *Undergraduate Women's Studies Program* by Carolyn Rhodes, Old Dominion University Research Foundation, November 1978, p. 58).

The following objectives could be derived from the goals listed above.

Faculty shall offer a course examining sex-role stereotypes and their social and cultural roots. The course shall include self-exploration and how stereotypes affect students and shall present means of overcoming them (process objective).

Students shall be able to respond to simulated real life situations in a non-stereotypic manner (impact objective).

Faculty shall develop and teach feminist focused courses in history, literature, sociology, and psychology (process objective).

Students shall be able to analyze content in sociology, history, literature, and psychology from a feminist perspective (impact objective).

Faculty shall develop and teach an interdisciplinary course that traces the development of women's roles in 19th century America (process objective).

Students shall be able to apply concepts from sociology, psychology, literature, and history in an essay about specific women in 19th century America (impact objective).

The goals and objectives cited above are all framed around desirable outcomes for students. Women's studies programs frequently also have objectives that concern faculty and the institution as a whole. For faculty, objectives may include acknowledgment of women's studies as a legitimate academic discipline, acceptance by faculty that students should be encouraged to take women's studies courses, and realization that a feminist perspective should permeate academic disciplines. For an institution as a whole, objectives could include coordination among the many offerings on and off campus that relate to women, such as women's centers or community organizations. Whatever the specific objectives of women's studies programs, an evaluation can address whether these objectives have been attained and how much progress has been made in attaining those objectives.

Before an evaluation can be designed to measure the achievement of objectives, a few additional steps are helpful. First, after stating all objectives, staff should categorize them. The categories should reflect the differences between process and impact objectives and also the target of the objective. For example, there may be objectives related to student learning, to student attitudes, and to the impact of the program on the total institution. Once categories are developed, staff should choose among those categories and objectives so that the program can be judged

in light of the most important issues. Generally, there are too many objectives in a program for an evaluation to be all-inclusive. The selection process, then, helps frame the design.

A Caveat

There may be a dilemma involved in designing an evaluation around explicit statements of goals and objectives. For women's studies, for example, the long-range goal may be to infuse the entire curriculum with feminism. As a path toward that goal, some intermediate goals—developing specific feminist courses—and objectives may be both necessary and desirable. Faculty and staff are more able to control and to measure the achievement of the latter goals, but, in the long run, they may be more concerned with the former. If the evaluation looks only at the feminist courses, it will not present guideposts or ideas for improving the more general atmosphere for women on campus.

Porter and Eileenchild define feminism as ". . . the desire to increase the power and autonomy of women as individuals and as a group (so that) they will be able to make informed and flexible choices in their education and their lives."[3] If a women's studies program accepts this definition and sees itself as both catalytic for others *and* important in its own right, it must take care to be evaluated in terms of both long- and short-range goals.

The problem, of course, is that the achievement of the longer term goals are seldom under the control of feminists themselves. Since evaluation results frequently are used to control programs, it is tempting to omit questions that are neither controllable by the program nor apt to lead to positive answers. On the other hand, omitting such questions deprives feminists of what may be important information. The nature of dilemmas is that there are no easy answers. Here, again, program developers and staff must decide what they want and indicate that to the evaluators.

---

[3]Nancy M. Porter and Margaret T. Eileenchild, *The Effectiveness of Women's Studies Teaching* (Washington, D.C.: National Institute of Education, in press), p. 2.

3

RESEARCH DESIGNS

Research designs are built around specific objectives. Often the most difficult step in constructing a research design is the first one—deciding what objectives the evaluation is going to address. An initial list of objectives for a program is often long, cumbersome, and totally unwieldy. It may contain 15 or 20 objectives. An initial list of objectives for women's studies programs could be equally long and unwieldy, especially when one considers objectives for students and graduates, for faculty and administration, and for the program and its impact on other efforts. The key, as we mentioned in the previous chapter, is to rank the objectives in order of priority on the basis of what objectives are the most important (and to whom) and what objectives must be addressed this year rather than next year.

Once a list of objectives has reached a manageable size, one examines each objective in turn to decide what evidence would be needed to ascertain the extent to which the objective was accomplished. The question to pose is what evidence would be the most convincing to show that the program is doing what one wants it to do.

This chapter contains two sections. The first takes a single objective and describes a design to measure its attainment. We then ask a series of questions about information not included in the design and build successively more complicated designs to incorporate that information. In practice, one tries to anticipate these questions in advance of any data collection or analysis. The second section, building upon the first, presents a complete hypothetical research design for the evaluation of a women's studies program.

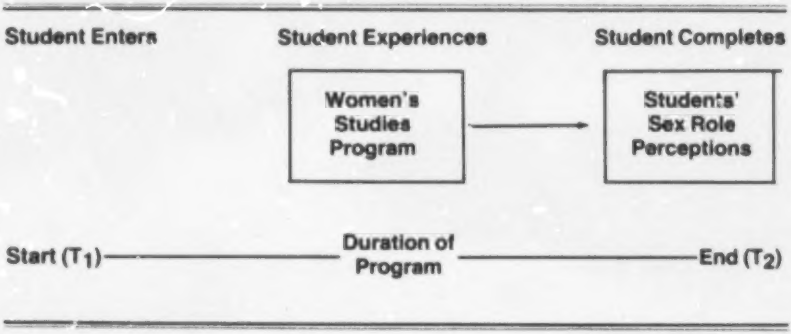
Design Strategies

Assume that the initial objective is the following: At the end of a one-year introductory course sequence in women's studies, students shall have equitable sex role perceptions of women's roles in society.

While the concept may be difficult to measure, the design strategy is

quite simple. Students would be asked at the end of the year about sex-stereotypic views, probably through a multiple-choice instrument on sex role perceptions (several instruments of this type are on the market). This design is called a *post-only design*, since students are queried only at the end of a program and only students in the program are asked questions.

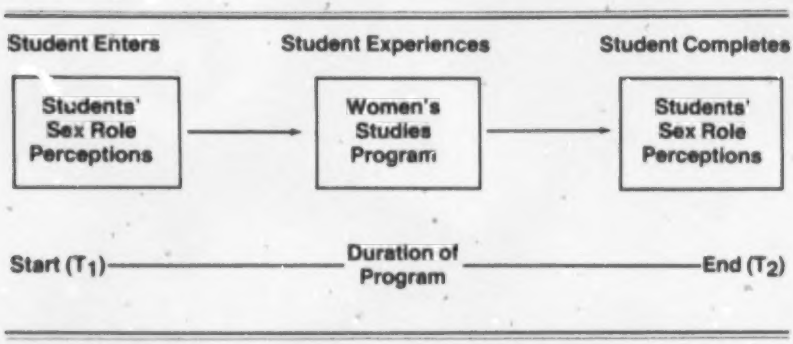
Schematically, the post-only design looks like the following.



Someone, however, may ask: "Well, the students in the program probably didn't have sex-stereotypic views when they went into the program anyway, so how did the women's studies program make a difference?"

To respond to that question, a somewhat different design is needed. Students would be asked a series of questions at the beginning of the program prior to any instruction and then again at the end of the program. The two sets of responses then would be compared to see whether there were differences in their scores. This is called a *pre-post design* and involves only the students in the program.

Schematically, the pre-post design looks like the following.

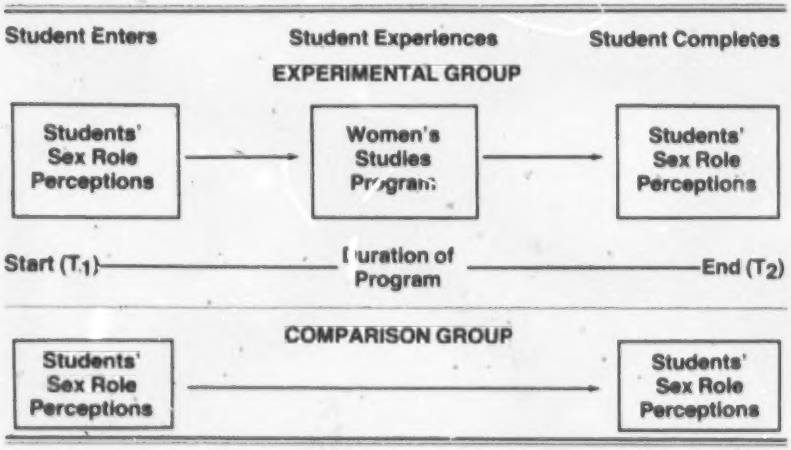


But once again, someone may ask: "It's fine that we know the students now have less stereotypic views of women's and men's roles than when they first enrolled in the women's studies program. But how do we know that what they learned couldn't have been learned just by being on campus, or from their friends, or from the women's movement in general?"

To respond to this question, a still more complex design is required, since we now must involve other students, ones not in the women's studies program. We need not only the pre-post design (for the beginning and end of the year), but also some kind of control or comparison group. A control group would consist of students who are identical in all respects to the students in the women's studies program but are not enrolled in it. In field experiments, a control group is usually obtained by overrecruiting for a program and then randomly assigning students from the pool to the program (experimental group) and to the control group. The research design that uses experimental and control groups is called an *experimental design*.

Since control groups are almost impossible to isolate in real life, most evaluation work these days uses comparison groups. Examples of comparison groups for women's studies may be a random sample of students from the rest of the college or university or a random sample of students within a given department (in the event that the women's studies program was only open to students within a given department). The research design that uses experimental (women's studies students) and comparison groups is called a *quasi-experimental design*. These are the strongest yet most realistic designs since they enable us not only to examine program effects over time but also to determine whether any changes are the result of the program itself or stem from other factors.

Schematically, the quasi-experimental design looks like the following.



So far, we have discussed only one objective related to student learning. We have not yet addressed the multiple objectives of women's studies programs for students, faculty, and the institution as a whole. The next section proposes a sample evaluation design that has multiple objectives.

A Sample Comprehensive Research Design

Let's assume that the objectives of a women's studies program have been defined and listed in order of priority. Assume that there are three objectives for students in the women's studies program, one program objective for graduates, one objective for department faculty, and one administrative objective for coordinated work with a campus-based women's center.

A description of each of the objectives and of the composite research design follow. It should be noted that the objectives have been selected more for their diversity than for their cohesiveness. One could say that the first three objectives are of prime importance to the faculty teaching the specific courses, that the next three objectives are of prime concern to the administrative staff in the college or university, and that the final objective is a first step in the long range plan of the women's studies coordinator to infuse feminism throughout the campus.

Over the course of the year, students in the women's studies program will have less stereotypic views of women's roles in society; these views can be attributed to the program and not to other factors (impact objective).

This example was detailed in the preceding section. In short, at the beginning of the year, students in the women's studies program and a comparison group of students randomly selected from the college or university as a whole or from the department in which the course is housed shall be given (under identical administrative conditions) an instrument on sex role perceptions. The same instrument shall be re-administered to the same students at the end of the year. Changes in scores from the beginning and end of the year shall be compared for each group, as will differences in scores between the two groups.

Students in women's studies programs shall have developed leadership skills and shall feel more confident in making presentations before groups (impact objective).

This is really two objectives, one related to developing leadership skills and one related to self-confidence. They should be discussed separately and are combined here only as a matter of convenience for the discussion. Changes in leadership skills and self-confidence are expected over time, so there is need for a pre-post design, using only the women's studies students. One could measure the attainment of this objective in a

number of ways, including having a panel of consultants rate students on various points. These consultants would look at how the students present themselves before groups at the beginning and end of the year, observe at different points in time how students work in small groups to organize tasks and to carry them out. They might ask the students themselves whether they feel, over the course of the year, they have developed leadership skills and a greater sense of self-confidence and self-assurance in what they do. Other ways of collecting information over two or more time points include using peer review, video tapes, and self-esteem rating scales.

Students in the women's studies course on women in the community shall prepare resource materials on services available to women in the community, such as women's centers, rape crisis centers, CETA projects, abortion clinics, social security offices, employment agencies, and child care facilities, and shall outline where such services are located, how they are staffed, what qualifications exist for being eligible for services, how much services cost, and the name and phone number of someone to call for further information (impact objective).

This is a relatively simple objective to measure. One need look only at whether the objective was accomplished (a post-only design). The main questions to ask in measuring the objective are whether all services have been described, whether the information is complete and has been double-checked, and whether the information has been shared with and understood by the other members of the class. One could question, however, whether this should be the complete objective. Once this set of useful resource material has been compiled, one could argue that the objective should also include some activities to make these resources known and available to the community at large.

Over the course of the academic year, there will be a growing awareness among students on campus of the women's studies program, and more students shall plan on enrolling in women's studies courses during the following year (impact objective).

This is also an objective that requires information to be collected at two different points in time so that one can see whether more students become aware of women's studies programs and want to enroll in them. Information could be collected through a brief postcard survey of a random sample of students at the beginning of the year and of another random sample of students at the end of the year. Such surveys might ask whether they were aware of the women's studies program (prior to receiving the postcard describing it) and whether they are interested in enrolling in courses (course offerings could be listed on the postcard). The number of responses would then be tallied and changes in percentages from the beginning to the end of the year could be compared.

Graduates of the interdisciplinary women's studies program shall assess the program as very useful in their further education or in employment (impact objective).

While the data collection design for this objective is quite straightforward (a one-time only administration, usually of a mailed survey), both getting the data and interpreting the results are problematic. Among the first matters to decide are: how long a person must be a graduate of the program to be included in the sample; how to find addresses for these graduates so they will receive the survey; how to get a decent response rate (that is, how to make sure most of the people return the survey form). One could decide, for example, to send surveys only to those graduates who will have been out of the program one year. One then could make sure forwarding addresses are collected before the students graduated and could inform students that a survey of graduates would be undertaken, stressing how important it is for the program to find out what each of them is doing one year hence. In fact, if the last two steps are not taken well in advance, it may not be possible to do a study of graduates at all. Within the mail survey or telephone interview or both, one could ask what the graduates currently are doing and whether they find the courses they took in women's studies useful in what they are doing. One could also ask for examples of how the program has been useful, both professionally and personally. It also may be helpful to get anecdotal data—specific information about specific graduates—to see whether these data match the more general information obtained through the survey items. While one cannot generalize from the anecdotal information, the use of individual cases and short success stories can be very powerful.

What is most problematic about measuring the attainment of this objective is that pursuing further education and obtaining employment involve many factors that are unrelated to the degree in women's studies. It is also the case that it may take 5 to 10 years to move into a good professional position. Furthermore, if one hasn't the financial resources for additional education or if one cannot find a job because of tight labor markets, one cannot fault the women's studies programs, even though some may try to do that.

The problems associated with measuring the objective should not, however, preclude women's studies directors from conducting such surveys (we will be talking more about surveys in the next chapter). Alumni offices or career placement offices within the college or university may cooperate in doing such a study and may help with the mechanics of such a survey. Little is known of the activities of women's studies graduates, and a continuing concern of many, both within and outside the women's studies field, is that the degree hampers, or at least doesn't support, career options for graduates. Without the data, one cannot address these concerns.

By the end of the academic year, the majority of faculty in the department of history shall recommend to the students they counsel that the students enroll in the women's studies course on women in 19th century American life and shall recommend the course be required for a departmental major (impact objective).

This is another case where two objectives were combined into one. In constructing a research design, one should treat them separately. These objectives are not as simple to measure as they seem. They may involve asking the history faculty whether they are recommending taking the particular course and whether there are plans afoot to have the course become a required course for a departmental major, but one also may want to ask students their perceptions to see if faculty counselor behavior is as described. Often it isn't. The objectives are included here not because they are relatively easy to measure, but because they allow us to raise a sign of caution. These are objectives over which the women's studies program may have little control or influence. In deciding whether to include them as objectives, program staff should consider whether they relate to their long-term goals and what they will do with the information once they have it.

The women's studies coordinator and the director of the campus-based women's center shall develop and implement a career and group counseling program for women students at the university (process objective).

This is a process objective; we are concerned with how the career and group counseling program develops, who is involved with its development, and what factors seem to have helped or impeded the development of the program (including costs, staff time, space). Information about how the activity is proceeding is usually collected over several points in time in order to assist the developers in creating and managing the activity. Such information also becomes useful when other activities of this type are planned, since this type of evaluation (formative evaluation) often identifies roadblocks that must be overcome.

Having discussed each of the seven objectives with designs for measuring their attainment, we now can summarize the overall design for the evaluation of the program. The summary is presented in Figure 1 to assist program directors not only in planning data collection and analysis but also in deciding whether one has the resources (both staff time and funds) to conduct the evaluation. Just as there is no single type of evaluation, there is no single evaluation design. Depending upon the information needs of various audiences and the resources available, evaluation designs can be either expanded or restricted. What is crucial about developing the evaluation design is the knowledge it provides women's studies directors *prior* to making commitments about what evaluation activities will be undertaken.

The following chapters provide additional information about the pros and cons of various methods of data collection, data analysis, and reporting.

Figure 1

SAMPLE EVALUATION PLAN*

OBJECTIVE AND RESPONDENT GROUP	TIME SPAN		
	Before or As Program Gets Underway	While Program Underway	Near or at End of the Program
Sex Role Perceptions			
Women's studies students	Sex role perceptions (instrument)		Sex role perceptions (instrument)
Comparison group	Sex role perceptions (instrument)		Sex role perceptions (instrument)
Leadership Skills			
Women's studies students	Observation before groups (consultants)	Observation before groups (consultants)	Observation before groups (consultants)
	Self-report on self-confidence (survey/interview)		Self-report on self-confidence (survey/interview)
Compiling Community Materials			
Women's studies students			Resource materials produced (faculty review)

*Method of data collection indicated in parentheses.

OBJECTIVE AND RESPONDENT GROUP	TIME SPAN		
	Before or As Program Gets Underway	While Program Underway	Near or at End of the Program
Awareness and Interest in Women's Studies Comparison groups (one at beginning, another at end of year)	Aware of/interested in taking women's studies (postcard)		Aware of/interested in taking women's studies (postcard)
Impact of Women's Studies Degree Graduates of women's studies		Perceived usefulness of degree, current activities (mail survey)	
Impact of Women's Studies on a Department All history department faculty			Students counseled to take women's studies, course becoming a requirement of department major (interview)
Increasing Institutional Impact (Career and Group Counseling Program) Women's studies director and women's center director	Informal interviews usually tied to the specific calendar of activity about how the counseling program is developing and how roadblocks to its establishment are overcome.		

4

INSTRUMENTS AND ITEMS

In discussions of what and how important specific objectives are, it is useful to ask, "How are we going to know whether we have achieved them? What information would we and others consider believable and convincing?" The data collection instruments provide that information. This chapter is designed to assist directors of women's studies programs in reviewing instruments and in knowing the questions to ask about instruments rather than to teach them to construct instruments themselves. We shall provide some sources about both available instruments and instrument construction.

A variety of data collection instruments can be used (from multiple choice surveys to structured interviews to observation). Each has its own advantages and disadvantages. In this chapter we shall be discussing various forms of data collection, concentrating on what instruments are most appropriate to use for different objectives. Before we get into a discussion of instruments, however, several tips on instruments in general may be useful.

One of the first questions to ask is, "What do we really need to know?" One of the most common errors made in collecting data is using instruments that ask for information because everyone asks for it; too many instruments include items that never find their way into any data analysis.

One may at some point, for example, want to know what students' future plans are, whether they have decided upon major fields, whether they have scholarships or some other form of financial aid, what their marital status is, and whether they have done community organizing; but unless one has decided to use these data in the analysis (and has objectives requiring that kind of information), one must question their appearance on an instrument. Throughout any evaluation, one must continue to ask, "Am I going to use this information to improve the program? Is this information going to be useful to people who are making decisions about the women's studies program?" Unless one has strong feelings that the information is going to be useful, it may be wiser not to

include it. The addition of what may be superfluous items will only take away time from more important questions, and can lead to disgruntled respondents.

How one asks for information (the method of data collection) is as important as the content of the specific questions. The more impressionistic or affective the information sought, the more likely it should be collected face-to-face rather than through a survey. If an objective relates to changes in behavior, it is preferable to use instruments (e.g., observational techniques) where one can in fact examine behavior rather than ask respondents about it.

Before setting about designing instruments and items for a women's studies program, look at what instruments are already available. The creation of instruments, whether they be surveys or interviews or checklists, appears disarmingly simple, when in fact it is a craft (and some would even say an art form) requiring considerable precision.

For example, assume one wants to find out whether students found a program or course valuable or useful. One could ask, "Did you find the women's studies program valuable? Answer yes or no." Regardless of what the distribution of responses is (from 100 percent yes to 100 percent no), how is one going to be able to interpret the data? Someone could well ask, "Valuable for what? Compared to what? In what context? What was valuable—the content, the teaching style, the contact with other students?" Unless the instrument is designed to probe these other questions, it is unlikely one will learn anything about the true value of the program. As a guide, each question should seek one defined piece of information about one single item.

There are also pitfalls to be avoided with a completely structured item with multiple choice answers. Take, for example, the following item:

> I have found the introductory women's studies course useful in:
> 1. Deciding upon my major and in making my career plans;
> 2. Finding a support group within the college/university;
> 3. Providing me with examples of the contributions women have made in the social sciences;
> 4. Making explicit the values imbedded in what is included in and excluded from traditional courses.

Problems with the item include:

- Not telling the respondent either to check all that apply or to mark the most important item;
- Giving the respondent no opportunity to say that the course was not useful;
- Combining two or more answers in one response (in the first response, one or the other or both may be true, but the evaluator can't tell which); and

- Having a global comparison that may not be true (those who mark the fourth response may not have a common frame of reference about what traditional courses are).

Solutions might involve asking respondents to rank responses in order of importance, breaking the questions down into several smaller questions, permitting another response category, clarifying the broadly defined terms, and/or scaling responses from positive to negative.

One can never construct the perfect instrument the first time around, nor is it often possible to judge in advance whether some items will be easily understood by the respondents. Consequently, the use of pilot tests is very important. Pretesting or piloting an instrument can occur in several ways. The most common form is to ask a group of respondents (similar in background and experience to the respondents one wants in the study, but perferably not those one will query later) to fill out the survey form or complete the interview. This is done in part to get responses to the items, but mostly to find out whether the questions are clear and sensible and whether the possible responses (on multiple choice items) are appropriate and complete. Such pilot tests also are used to find out how much time is needed to complete the instrument. The responses and comments then serve as the basis for revisions.

Survey Instruments

Survey instruments usually consist of multiple choice items, rating scales (the familiar five-point scales of "always" to "never," or "strongly agree" to "strongly disagree"), or checklists. They may be packaged, standardized instruments or ones constructed by the program staff or evaluators. They are very simple to administer and analyze (especially by computer), require little time to complete, and provide the least amount of information about a program or course. Since all the answers are provided (respondents check just whatever applies), one can only learn what one already has asked.

Standardized instruments (available through source books) save development time and money, which can be considerable. Usually, though not always, information is provided about how good the instruments are, whether they are valid (measure what they say they measure) and reliable (measure consistently), what the intended respondent groups are, and what concepts they are trying to measure. The main concern with standardized instruments is that they do not necessarily relate to what is taught in a program or course. Consequently, one should examine readymade instruments very carefully to see if they can apply to one's own program objectives.

One of the mystiques about standardized tests, such as norm-referenced achievement tests, is their aura of eternal truth, as if the tests measured some universal. It is easy to slip into giving them more credibility than they warrant. For example, in evaluating student achieve-

ment in an alternative high school, questions were raised about whether the tests would be appropriate considering the school's nontraditional curriculum. When the students' scores were higher than expected and their gains over the year much greater than those for a comparison group, however, the questions of appropriateness were dropped, to be replaced with statements about program success.

American Institutes for Research, under contract to the Women's Educational Equity Act Program, has prepared a *Sourcebook on Measures of Women's Educational Equity* containing a number of survey instruments. Since many project directors and others spend a great deal of time and effort in developing measurement instruments to evaluate the progress and outcomes of their programs, this collection should ease the burden on project directors and on directors of women's studies programs.

Some 167 instruments are described in the *Sourcebook*, more than 60 of which may be helpful in the evaluation of women's studies programs. The description of each instrument includes measurement variables, target population, instrument type, administrative procedures, scoring, description of sample items, availability of instrument, and references.

Measures that may be of use in women's studies programs include attitudes toward women's studies as an academic discipline, achievement motivation, fear of success, sex role stereotyping in careers, attitudes toward feminist issues, sex role perceptions, Title IX, career awareness, awareness of sexism, and attitudes toward mathematics.

In the winter of 1979, the *Sourcebook* will be available through the Education Development Center, Inc., 55 Chapel Street, Newton, Massachusetts 02160.

Also developed by American Institutes for Research is *Measures of Educational Equity for Women: A Research Monograph.* It too contains descriptions of a number of instruments useful to directors. It can be obtained through American Institutes for Research, P.O. Box 1113, Palo Alto, California 94302.

Interviews

Interviews are most often used when: (1) one doesn't really know the range or content of all possible answers; (2) the answers may be somewhat sensitive, and respondents are concerned about using their own particular wording in responding; and (3) the follow-up or probing questions depend upon responses that—because of the complexity of the issues addressed—cannot be predicted accurately in advance.

Interviews may be structured or unstructured. Structured interviews have pre-established questions in an ordered sequence. Responses may be open-ended or solicited to fall into predetermined response categories. This type of interview may be the most useful when analysis time is

limited, when many different interviewers are used, or when experience with interviewing is limited. Structured interviews are more often used where there is no face-to-face contact (such as in telephone interviews) and when the interviewer is not particularly knowledgeable about the program. Unstructured interviews (which customarily have very general questions with a variety of probes based on initial responses) are more commonly used in face-to-face situations, when the interviewer is experienced with the interviewing process and the program, when few interviewers are used, or when much time is available for careful analysis of data to draw out common patterns of response.

The format of interview questions is quite different from multiple choice items, although some multiple choice items may be used in an interview setting, provided the form is given to the respondent to complete. The interview setting cannot be used to ask respondents to do such things as "rank the following 10 items in order of importance," for respondents usually can neither retain nor reorganize that much information without being able to see the items written down. On the other hand, if one does not have pre-set responses, the interview is quite useful in probing what the respondent feels is most important. Unlike a survey form, the interview is active rather than passive; the responses are generated by the respondent rather than from among a predetermined list.

There are both advantages and disadvantages to the interview approach. The advantages include freedom to respond without constraint and the opportunity to elicit unanticipated information. An interview puts a greater burden on both the respondent and the interviewer, since there are no pre-set answers and the process is usually more time consuming. It overcomes the biases the respondent may have against written forms, but assumes that respondents are verbal and articulate. The burden on the interviewer is greater as well. The quality of the interview depends upon how good the interviewer is in establishing a sense of rapport and trust, and in recording all relevant information. Interviews are more difficult to analyze unless the questions, however general, are framed around the main issues the evaluation is to address, which once again brings us back to the research design. As with survey instruments, interview items should focus on those objectives that are most important to the evaluation.

An excellent, detailed guide to all aspects of interviewing is contained in the *Interviewer's Manual* put out by the Survey Research Center, Institute for Social Research, the University of Michigan, Ann Arbor. Revised in 1976 and available in a paperbound edition, it addresses many of the issues in interviewing and includes many helpful tips on the conduct of interviews.

Other valuable resources for interviews about women's studies programs are those people who have been involved in reviews and evaluations of women's studies programs. Nancy Porter addresses some of these issues in an article, "Evaluation: Reflections of a Program Con-

sultant," in the Fall 1977 issue of the *Women's Studies Newsletter*.

Florence Howe has made a number of valuable suggestions about interviewing based upon her field work for *Seven Years Later*, a book commissioned by the National Advisory Council on Women's Educational Programs. She suggests that it is important to interview as many people outside of a program as inside, asking the same questions of teachers, students, and administrators, including coordinators or directors of ethnic studies and other inter-disciplinary programs to see how the institution treats all deviations from traditional programs. During the conduct of the interview, she found it useful to have several questions for each interview that the interviewer knows routinely and that move from eliciting information about the person (e.g., title and function at this college, formal or informal relationship to the women's studies program) to general information about the program itself (e.g., its strengths and weaknesses, how the person would like and not like to see the program develop during the next five years). While large group interviews can be unwieldy, she found it also useful to have small group interviews with four or five people, since such interviews often sparked a discussion. Howe emphasizes one of the cardinal rules of interviewing: The interviewer is a collector of information and not a participant. The interviewer is not to bring her/his own ideas into an interview; the interviewer is always to remain a neutral party encouraging full and complete responses from the interviewees but is not to interject views about an issue or program. Lastly, Howe noted that the most difficult and yet essential information to gather about a program is its history, not only to learn about the development of a program but also to understand the social and political dynamics of the individuals in the program. If at all possible, this information should be obtained in advance or at the beginning of the data collection so that it forms a base against which to measure other responses and aids in the development of probing questions.

As is also the case with surveys, it is often difficult to know in advance all of the questions that should be asked in an interview setting. Thus, some pilot testing or pretesting of the interview schedule is recommended, not only to clarify questions and estimate response time but also to serve as a device for interviewer training.

Observation

This method of data collection derives from the ethnographic techniques of anthropology and can be very helpful in those situations where one is concerned more with behavioral change than with affective or cognitive learning. It is also very useful in looking at unanticipated consequences, since observation can be used without having a predefined range of actions in mind. In those cases where observation is being used to see whether a specific behavioral objective has been accomplished, one

needs to ask whether the observers are observing the same thing and whether they have been given adequate instructions about what they are to look for. We all have selective perceptions, so what may be important to one person may be missed by another. The greatest strengths of observational techniques are also their greatest weaknesses. While an observer provides an independent, detached view, one must try to make sure that observers focus on the phenomena of importance. Observers should be well briefed on what one is trying to accomplish. In this case, checklists of items to look for can be very helpful to observers, especially if they are then encouraged to explore other areas as well.

For example, if two objectives are to have students develop better leadership skills and to become more self-confident in making presentations before groups, then observers should know what leadership skills the program director has in mind and what signs of self-confidence seem appropriate. Suppose a class is divided into small working groups with instructions to complete a particular project. An observer could then see: (1) how the leader is chosen; (2) how the leader behaves; and (3) how tasks are allotted to group members. Since the objective has an explicit time sequence, observers would be used at several different points to track development. In examining self-confidence, observers would concentrate as much on presentation of self as on content of materials presented.

The ethnographic literature, including the growing field of anthropology in education, could provide to women's studies directors some useful guides to observer selection and training. As with other approaches to data collection, there is no single approach to observation, and the literature is abundant with observational approaches tailor-made to particular situations and objectives. Stephen Wilson's article on "The Use of Ethnographic Techniques in Educational Research," published in the Spring 1977 issue of the *Review of Educational Research*, is a very good introductory piece on the rationale behind and processes of this approach.

What is most important to remember about instruments—about the methods of data collection—is that they are *the* source of information about whether one has accomplished what one is trying to do. Instruments must be carefully examined to see whether they address all the areas of interest, include all the groups one is concerned with (e.g., students, faculty, administrators), and are technically sound. In addition, instruments (from surveys to observation guides) should not be administered until one has a good idea about how one is going to analyze and present the information derived from them. The next chapter addresses the major areas of data analysis.

5

DATA ANALYSIS

As with the chapter on instruments, the purpose of this chapter is to acquaint women's studies directors with the different types of data analysis, primarily so that they may become better equipped to discuss analysis issues with others.

Data analysis is concerned with reducing the bulk of information obtained through all the instruments into manageable form. It allows one to look for patterns within a group of respondents and across groups of respondents and, with some kinds of data, to determine whether significant relationships exist.

A data analysis plan should be developed before any data collection takes place. While it is not always possible to be precise about all the analyses to be performed, it is crucial to have a general plan of analysis tied to the objectives, and to know how much time and money are available for analysis. Access to any computer services also must be determined in advance.

Whatever the research design selected and the kinds of instruments used, a variety of data analysis techniques exist. Data analysis is often the most baffling part of an evaluation, in part because theories of probability and statistical vocabularies are introduced. There is, however, nothing mystical about data analysis. Every data analysis technique has an English language equivalent.

In fact, much of the data analysis appropriate to evaluating innovative programs, including women's studies, does not involve the use of the more complicated statistical techniques, since the assumptions upon which these statistics rest are often not relevant to real world programs.

On the following pages, the range of data analysis possibilities are illustrated. One objective, for example, may be to tie the women's studies program more closely to some of the institution's goals. In discussions with administrators, one may have learned that the institution is interested in increasing the enrollments of nontraditional students (such as women in their midthirties who are returning to college after their children are in school). One may find that, from 1977 to 1979, enrollments of these women in the college/university increased from 5 to 10 percent

of the total enrollment. By looking at the enrollment of students in women's studies programs over the same period, one may find the enrollment of these women increased from 10 to 30 percent (that is, 30 percent of these women were in women's studies courses in 1979). Stated another way, while the enrollment of returning women doubled in the past three years, their enrollment in women's studies programs tripled. One also may find, through a postcard survey of students throughout the college or university, that half of the returning women are very interested in taking a women's studies course. By using these percentages—more formally called descriptive statistics—one can present a case to the administration that the women's studies program is an important attraction to these students and assists the institution's goals.

With a number of the sample objectives used in the preceding chapters, the data analysis relies on descriptive statistics (such as percentages). These objectives include students' growing awareness of the women's studies program, an increase in the number of students planning to enroll in women's studies courses the following year, and the majority of faculty in a department recommending students enroll in a women's studies course. Data from graduates on the usefulness of the program in their further education or employment also can be analyzed using descriptive statistics

Much interview and observational data are analyzed using much the same approach as a journalistic investigator. Assume that one objective is to have a course on women in English literature become a required course for a departmental major. Part of the analysis regarding the extent to which the objective has been reached could read as follows.

> Half of the faculty in the English department, the faculty person teaching the course on women in English literature, the chair of the department, the women's studies coordinator, and the dean of the college were all interviewed separately about the prospects of the course becoming a required course for a departmental major. While the coordinator and faculty member teaching the course felt the course would complement the existing requirements, the faculty queried thought it superfluous. Yet when asked whether they had reviewed the reading list, talked with the faculty member teaching the course, and discussed the course content with any students, all responded no. The chair, with the concurrence of the dean, has secured stable funding for the course but was reluctant to pursue requiring it, fearing it would lead to a proliferation of required courses.

This excerpt is essentially a condensation of a series of interviews with concerned parties obtained through a structured interview format followed by some probing questions. There is nothing quantitative about

the analysis, but it is data analysis nonetheless.

The more quantitative data analysis usually appears in efforts to assess changes in student attitudes or learning. The following objective, cited in Chapter 3, for example, requires statistical analysis.

> Over the course of the year, students in the women's studies program will have less stereotypic views of women's roles in society, and these views can be attributed to the program and not to other factors.

Here one would probably use one of the available standardized instruments in which numerical scores are derived from a set of items using a "strongly agree" to "strongly disagree" scale. Statistical techniques then would be used to look at the distribution of scores to see whether differences among the scores for the two time periods (pre and post) and the two groups (women's studies students and the comparison group) could be based on chance variation or whether the scores follow a sufficiently set pattern so that differences are most probably not due to chance.

For specifics on what techniques should be used in what situations, refer to materials listed in Appendix C: Resource Materials. Two recent publications may be quite useful since they present step-by-step statistical analysis. The *WEEA Evaluation Handbook*, prepared by the American Institutes for Research, is designed for individuals who want to submit proposals to the Women's Educational Equity Act and need help in developing adequate evaluation plans, and for those who already have projects but need help in conducting evaluations. While the *Handbook* does not specifically address goals and objectives of women's studies programs, the chapter on input assessment (how questions affect the designs of evaluation) and the chapter on statistical methods (presented in a step-by-step fashion) may be quite helpful. The eight-volume *Program Evaluation Kit* by Lynn Lyons Morris, Carol Taylor Fitz-Gibbon, and Marlene E. Henerson—although geared for use in school districts—may be helpful as well. The statistical analysis volume is one of the clearest to follow. The kit, in single volumes or as a complete set, is available from Sage Publications, Inc., P.O. Box 5024, Beverly Hills, California 90210. Costs of single volumes range from $4.50 to $8.50. Other references to data analysis techniques are included in Appendix C.

6

REPORTING RESULTS

In the first chapter, we indicated that reporting the results of the evaluation is an integral aspect of the job. There are two issues involved in providing evaluation feedback: To whom should it be given? How should it be presented?

Audiences for Evaluation

If an outside evaluator is hired to assess a women's studies program, that person has an ethical responsibility to report results to whomever did the hiring.[1] The client is the first audience, and that is not always the women's studies program staff. It may be a dean or other administrator. In those cases in which someone has commissioned an evaluation with the intent of killing the program, there is really no way program staff can avoid that person's getting the results first. They can, however, try to be simultaneous recipients of the results so they are able to present their own interpretations. Because the evaluator owes the client the first hearing, it is probably to the program's advantage to be the client.

Besides the legal client, others should be informed of the evaluation results. Anyone who participated in the evaluation has a right to some feedback about its results. Students and potential students have an interest in what is found about a program, as do staff and faculty in related fields. College and university administrators are another audience for the results. In addition, if the program was developed with Federal, State, or foundation funds, these agencies are especially interested in the results of an evaluation even if it is not (and it frequently is) mandated. Finally, program staff and evaluators might well want to share evaluation experiences, methods and findings with others involved with women's studies so that knowledge can accumulate across programs. This process could involve reports (essays or brief research notes) to the *Women's Studies Newsletter* (Box 334, Old Westbury, New York 11568), work with regional affiliates of NWSA, and presentations at national conferences.

[1] Joint Committee on Evaluation Standards, *Standards for Education Evaluation*. Prepared for the Lily Foundation, the National Science Foundation, and the National Institute of Education, 1978, mimeographed, p. 101.

Copies of evaluation materials and findings sent to the NWSA national office would become available to others in the field.[2] We also would like to receive materials so we can revise and expand this handbook.

The lessons learned from Marcia Guttentag's pioneering work in the decision-theoretic approach to evaluation apply in any discussion of audiences. The more decisionmakers (regardless of funding source) are involved in setting the questions the evaluation is to address and in following the ongoing evaluation activities, the greater the likelihood the results will be used. While there are dilemmas here (e.g., what happens if we involve everybody and the results are all awful—is it curtains for women's studies?), in the vast majority of cases, it is always better to share some of the design responsibility with others.

Presenting Results

Results can be reported in many ways, ranging from a full-blown monograph with statistical data and computer printouts to an interactive seminar. What is important is to present the results in ways that meet the needs of the various audiences. Given the numbers and roles of people who might be interested, developing different modes of presentations for each might prove prohibitively expensive. In that case, it might be good to develop three products—the full report, an executive summary, and a press release—and augment those with whatever special materials are affordable.

No matter how the results are reported, the following information should be included:

- A program description (this can be very abbreviated if the audience consists solely of people who know the program well);
- The evaluation design and analytic techniques (written in nontechnical language for nontechnical audiences); and
- The results.

Further, whether presented in writing or orally, the report should be in clear, jargon-free English. Clarity of expression is a sign of clarity of thought, and jargon usually muddies issues rather than clarifying them. As a physician warned a friend about other doctors, "If they can't tell you what's wrong in plain English, they probably don't understand themselves."

In Appendix A is an example of an executive summary of an evaluation. It presents and answers program managers' questions about a number of women's studies and other programs in a direct and concise way.

Within this handbook, we have sought to provide information on evaluation of women's studies in a clear, concise manner. Two other appendixes, Abstracts of Studies and Resource Materials, report on results of previous work.

[2] The address for the National Women's Studies Association is: National Women's Studies Association, University of Maryland/College Park, College Park, Maryland 20742.

APPENDIX A: EXAMPLE OF EXECUTIVE SUMMARY

PROJECT WELD
WOMEN'S EDUCATION: LEARNING AND DOING

Funded by: The Fund for the Improvement of Postsecondary Education, the Department of Health, Education, and Welfare

Project Director: Margaret A. Talburtt,
Formative Evaluation Research Associates (FERA)
216 East Huron Street
Ann Arbor, Michigan 48104

Introduction

This project was undertaken to provide information about three types of educational options available to undergraduate women: internships, women's studies classes, and skills development classes (or workshops). An internship was defined as a college-sponsored program which had a stated goal of career exposure in an out-of-classroom setting, such as a summer job with a federal agency. A women's studies experience was defined as an academic course which focused on women's contributions or relationships to a particular field, such as feminist politics or women in history. Courses included in this group were *not* necessarily part of a formal women's studies program. The third category, skills development, signified a class or workshop which held the stated goal of behavioral change. Included were assertiveness training classes, career planning seminars, and personal growth classes. While these three program types did not exhaust the many options that have recently been designed to foster undergraduates' self-awareness, personal growth, and professional potential, these programs represented some of the more pervasive and successful of the recent innovations. The goal of this study was to evaluate the outcomes of these programs.

Method

Within this framework, eight schools were chosen for participation in this study. Six of the eight schools were predominantly women's colleges. The other two were women's centers at large coeducational universities. The rationale behind this choice was to begin an assessment of women's educational progress in settings which had committed themselves almost exclusively to the full development of women students. The eight schools represent a variety of institutions for women and formed a broad sample from which conclusions about women's education could be drawn. The participating schools or programs were: Wellesley College, Mt. Holyoke College, Cedar Crest College, Wells College, Westbrook College, Mt. Vernon College, Everywoman's Center at the University of Massachusetts, and The Women's Center at the University of Minnesota.

The study began with preliminary visits to the programs at each school during the fall of 1975. Of the eight schools, six had internships, seven had women's studies courses, and three had skills classes. A total of 270 students, 50 faculty and administrators, and 25 clients (or intern employers) were personally interviewed during the winter of 1976, using questionnaires designed for this study. Students were randomly selected from lists of those who had completed their involvement with the program approximately one year prior to their interview. The majority of students were 19-23 years of age; fewer than 20 of the women were 30 years of age or older.

Information was analyzed during the summer of 1976 and two reports were sent to each program to contribute to their program improvement. In this way, some measure of formative evaluation was possible. The major cross-school report was completed during the winter of 1977.

In addition, an advisory board was formed to review project progress, stimulate questions for analysis, and suggest applications of the project's findings. Members were Dr. Elizabeth Tidball of George Washington University, Dr. Sherry Penney of Yale University, Dr. Mariam Chamberlain of the Ford Foundation, and Ms. Carol Stoel of the Fund for the Improvement of Postsecondary Education. Their ideas, as well as those identified by the program directors, and the FERA staff, led to analysis of the following points.

Why did students choose these experiences?

Summary: Students chose these various experiences to gain the general outcomes promised by each. Sixty-nine percent of the interns (n = 128) sought career exposure, while 66 percent of the women's studies students (n = 82) desired new academic perspectives about women. Forty-four percent of the students in skills classes (n = 55) wanted new skills, such

as assertiveness or communication skills, but another 40 percent mentioned new ideas or new attitudes as additional expectations.

What did these various programs offer to students?

Summary: Students reported that the three types of programs achieved the goals that they promised. Internships promoted professional skills and career exposures. Women's studies classes fostered feminist perspectives and expanded concepts about what women can achieve. Skills classes led to increased self-understanding, assertiveness, and other skills related to personal effectiveness.

Quantitative student data gathered with standardized forms, however, also indicated that these various options achieved other important results which were *not* expected but were *highly* valued by students. Further, their programs accomplished most of these outcomes significantly better than the traditional (or general) curriculum at each school. For example, internships did more than provide career exposures. They increased self-confidence, openness, and assertiveness. Women's studies fostered self-confidence and a greater sense of personal potential. Skills classes affected self-confidence, independence, and feminist perspectives as well as interpersonal skills. In sum, each type of program accomplished *different* results, and all programs resulted in some outcomes that traditional educational experiences did not produce. Simply, innovative structures have impact on outcomes that students rate as important—they do make a difference, supplementing the traditional experience in some important ways.

What are the differences between single-sex colleges and coed universities?

Summary: Because this study did not include a representative or large sample of the varieties of especially coeducational schools, no firm conclusions can be drawn about differences among school types. It would appear, however, that special programs for women enhance the quality of education that a female undergraduate at a coeducational university receives, and *may* be even more important to her full development than similar programs at a women's college. Much more research needs to be conducted before such an inference can be accepted.

What is the impact of these programs?

Summary: All three types of programs had impact on undergraduates, and substantially more impact than other educational experiences.

Eighty-four percent of the interns and 85 percent of the students in skills classes said these experiences had "some" to "lots more" impact on them than their other courses and academic activities, whereas only 72 percent of the students in women's studies offered such ratings. Seemingly, the unique instructional processes and/or setting of the former explains the differential scoring. Nonetheless, the power of all three programs was obvious.

Were students satisfied with these experiences and what improvements would they offer?

Summary: Satisfaction with the process of the experience and with personal development was highest in the internships and skills classes, while satisfaction with academic content was slightly higher in the women's studies classes. Although this is not the way it has to be, it seemed that traditional instructional modes convey content better than innovative processes, while the nontraditional approaches seem to promote greater personal growth.

Despite the high satisfaction ratings, students did offer suggestions for improvement. In each type of experience, structure was the most often mentioned type of desired improvement. More academic content was the most frequently mentioned addition that students desired. There was some evidence that students may not know how to evaluate innovative educational experiences because they do not know exactly what to expect from them. Indeed, more than 75 percent of each group of students replied that positive unexpected outcomes had occurred in their respective programs.

Did these experiences relate to students' future roles?

Summary: Yes! Although a total of eight types of roles were mentioned, the major roles that students expected to assume are family and career. The most important skills they felt that they would need to be successful in those roles were self-confidence, interpersonal skills, and technical/professional skills. Clearly, these are the types of skills and attitudes that the three innovative educational experiences fostered. Within the groups, the interns were the most career oriented and the women's studies students the least family oriented. But to pretend that young women today do not seek family as well as career roles would be an erroneous assumption.

Did these students feel their institutions could be improved as schools for women?

Summary: Again, yes! Seventy-six percent of the women at single-sex schools and 57 percent of the women at coeducational universities felt that their colleges could be improved as schools for women. Academics, support services, and structural improvements were suggested as focuses for change.

Conclusion

This abstract contains many highlights of a full report. Specifically, useful data from the faculty interviews as well as more extensive student information are missing from this summary. The implications of these findings are many, however, and should be of interest to the students, faculty, program directors, and sponsors of programs for women. All of these groups should recognize the specific benefits of each type of program and relate these to their educational purposes in an integrated way. Clearly, each of these programs makes an important and unique contribution to the education of women and one that is not available in traditional higher educational settings.

APPENDIX B: ABSTRACTS OF STUDIES

*Documents with an ED number are contained in the Educational Resources Information Center (ERIC) collection. Microfiche or paper copies of these documents can be obtained for a fee from the ERIC Document Reproduction Service (EDRS), P.O. Box 190, Arlington, Virginia 22210.

Title: "Assessment of the Experiential Learning for College Credit in the Area of Women Studies"
Authors: Benjamin Sackmary and Hannah Hedrick
Source: Educational Resources Information Center, ED 155 208*

The paper proposes a general set of guidelines for granting experiential or life, learning credits. Examples from portfolios of women applying for women's studies credits based on their life experiences are presented, but there is no indication of whether—or how many—credits were granted. There are, however, samples of women's studies courses and goals from 31 colleges that may prove useful.

Title: "Attitudes of Students, Faculty, and Administrators Toward a University Women's Studies Program"
Authors: Glen L. DeBiasi and Carolyn Rhodes
Source: Old Dominion University, Norfolk, Virginia 23508

The study involved an investigation of attitudes toward a new women's studies program held by nonparticipant students, faculty, and administrators. A representative sample of each group was administered questionnaires concerning the awareness of and attitudes toward the program and its goals. Results showed that a relatively large percentage of the respondents were aware of the women's studies program, but fewer were aware of its specific goals. A majority were supportive of the program. A

relatively large percentage of faculty (60 percent) and administrators considered women's studies a valid field of academic inquiry.

Title: "Effects of Women's Studies Courses on Women's Attitudes and Goals"
Author: Eileen M. Canty
Source: Educational Resources Information Center, ED 150 490*

Canty used a pre-, post-design with a comparison group. The experimental group consisted of students in the Psychology of Women and Psychology of Adolescence courses; comparison group students were enrolled in Psychological Statistics. The measures were the Spence and Helmreich Attitudes Toward Women Scale, the California Psychological Inventory, Broverman's Sex Role Stereotype Questionnaire, and a detailed locally developed questionnaire. The results were that the experimental group had more liberal plans to combine marriage and career and to delay marriage. Further, there were changes in attitudes toward women's vocational, educational, and intellectual roles and toward marital relationships.

Title: "Evaluating Women's Studies: A Decision-Theoretic Approach
Authors: Marcia Guttentag, Lorelei R. Brush, Alice Ross Gold, Marnie W. Mueller, Sheila Tobias, and Marni Goldstein White
Source: *Signs*, Vol. 3, No. 4, 1978

The article presents an introduction to the Multi-Attribute Utility Decision-Theoretic (MAUT) model for conducting evaluation. The authors argue that the MAUT approach to evaluation is more effective for evaluating women's studies than traditional approaches for four reasons. First, the MAUT model addresses such questions as, "How can the program be improved?" and "How can it best meet the program goals of students, faculty, and administrators?" Second, the model allows different subgroups within a program to develop their own set of goals by which they can judge program success. Third, the MAUT model enables program personnel to change goals and/or means of achieving them at any time during the evaluation process. Finally, it allows for continuous evaluation of a program.

There are three major steps in the use of the MAUT model: "Stating the program's goals in the order of their priority and assessing the probability of achieving them, developing techniques to measure the

degree to which the program achieves those goals, and collecting and analyzing data generated by the administration of these measures."

The following steps are involved in stating priorities.

1. Identify all groups whose goals for the program should be evaluated.
2. Identify the issues to be addressed by the evaluation.
3. Identify the action alternatives to be evaluated.
4. Specify the goals of each group identified in the first step.
5. Rank the goals in priority order. Each constituency that has specified goals must agree on their importance.
6. Assign "importance weights" to each goal.
7. "Sum the importance weights, divide each . . . by the sum, and multiply by 100."
8. Judge the probability that each alternative can achieve each goal.
9. Calculate the total utility of each action alternative (the sum of the importance weight times the probability of success).

Measurement of goal achievement in the MAUT model conforms to normal standards of data gathering. Baysian statistics are used for data analyses.

Title: "An Evaluation of a Women's Studies Program"
Authors: Sharon A. Shueman and William E. Sedlacek
Source: *Journal of NAWDAC*, Fall 1977

The authors report on evaluation of the women's studies program at the University of Maryland, College Park. The evaluation emphasized input variables (the structure of the program and a description of the students) and output variables (short-term effects of the courses on attitudes and self-concepts). A comparison group was used. The measure included a questionnaire on student demographics, goals, and experience, one on student attitudes toward the course, the Bem Sex-Role Inventory, and the Tennessee Self-Concept Scale. The last two were used as pre- and post-tests. Significant changes were found on items that indicated awareness of sex roles but on no other items.

Title: "Final Evaluation of GCR 66: Images of Women in Literature, University of Minnesota"
Authors: Patricia Davis et al.
Source: Educational Resources Information Center, ED 124 240*

Davis and her associates recount the history of the course and its evaluation the third time it was offered. The history includes a discussion of the development of the syllabus as well as of the evaluation. Instructors

for the course were primarily responsible for the evaluation but received assistance from the University's Research Center.

In developing the evaluation plan, the instructors clarified their goals and objectives. The underlying evaluation questions were: (1) Did the course meet its objectives? and (2) Were the students different from a comparison group (those enrolled in "Man's Religious Beliefs?"). Using a pre-, post-design, students were tested on their knowledge of women in literature, the history of the women's movement, and the contributions of women to society as well as their attitudes toward women. The instruments testing the knowledge objectives were developed locally, and a standardized test was used to test attitudes toward women. In addition, an open-ended questionnaire was developed to ascertain students' expectations of the course.

The findings were generally positive and related to instruction.

Title: "Nursing: A Feminist Perspective"
Authors: Phyllis Kritek and Laurie Glass
Source: *Nursing Outlook*, March 1978

The authors recount the development of a course that was cross-listed in the school of nursing and the office of women's studies. The course was designed to bring a feminist perspective to bear on the nursing profession. Topics included: history of the women's movement as related to nursing; nurses as women (socialization, power, politics); sexism in health care; strategies for the future. The authors developed their own evaluation of the course. The evaluation included an assessment of student learning (examinations and evaluation of research projects), the Allport Study of Values, a self-developed evaluation tool, and student-written evaluations of films, lectures, discussions, and the instructors' performance. Results indicated that students valued the course, believed they learned something from it, and met their personal goals. Course evaluations were positive. The authors have used suggestions from the evaluations for modifying and improving the course.

Title: "The Paradox of Intention and Effect: A Women's Studies Course"
Authors: Lorelei R. Brush, Alice Ross Gold, and Marni Goldstein White
Source: *Signs*, Vol. 3, No. 4, 1978

The authors report a comprehensive, long-term evaluation of "the impact of an inter-disciplinary women's studies course on the self-concepts, sex-role attitudes, and sex-role stereotypes of the students

based on data gathered over a two-year period at a small liberal arts college." Two methods of data collection were used: an open-ended in-depth interview and pre-post administration of a questionnaire and of a battery of tests (e.g., Minnesota Women's scale, Broverman Sex-Role Stereotype Scale, the Bem Sex-Role Inventory). Students in other courses served as comparisons. The authors found that the in-depth interviews yielded more interesting and worthwhile data. Results indicate that while some students were deeply affected by the course, most changes were small and varied from year to year. Students did become more confident and able to defend their beliefs with information.

Title: "The Reported Impact of Women's Studies Courses on Students' Lives"
Authors: Allana Cummings Elovson and Irene Cockroft
Source: Educational Resources Information Center, ED 160 944*

Elovson and Cockroft describe an innovative approach to measuring the impact of women's studies courses on the lives of students—a type of goal-free evaluation. They developed an open-ended survey instrument with 23 questions that asked students to report the importance and impact of women's courses on their lives. The major innovation was the development of a scoring system that allows the coding of responses regardless of the specific questions that triggered them. Ninety-three percent of those surveyed reported positive impacts, including:

1. Increased confidence in and awareness of abilities;
2. Higher career aspirations;
3. Broadened perspectives on history and social issues;
4. Increased comprehension of political, economic, and social factors in shaping human experience;
5. Greater tolerance of those with different histories and value systems;
6. Increased desire to participate in and contribute to the study;
7. More questioning, analytical, and rational attitudes;
8. Less acceptance of exploitation and manipulation.

Title: *Seven Years Later: Women's Studies Programs in 1976*
Author: Florence Howe
Source: National Advisory Council on Women's Educational Programs, 1832 M Street, N.W., Washington, D.C. 20036

The review of women's studies programs was shaped by the following questions.

1. What is the current state of mature women's studies programs with respect to their faculties, students, curricula and classrooms, their

internal governance structures and administrative relationships to the universities in which they exist?
2. What has been the impact of these programs on their campuses and in their wider communities?
3. What dominant issues and needs are critical to the future of women's studies?

Howe found that enrollments in women's studies were growing, that a wide variety and depth of courses were being offered, and that there is a "growing realization of institutional impact."

Howe concludes:

- That the future of women's studies will depend both on the ability of its advocates to maintain their present high levels of energy, commitment and direction, and on the ability of the institutions to make the accommodations required not only by students, but also by faculty and programs;
- That while the energy commitment and direction of the advocates are abundantly present now, there are many who seek some stronger indicators that the productive and exciting level of women's studies teaching and scholarship is being provided or will be provided in the future, with commensurate institutional rewards.

Title: "Women in Continuing Education"
Authors: Linda M. Rosenwood and Patricia Lunneborg
Source: Educational Resources Information Center, ED 067 997*

The authors sought to find the effects of continuing education on women's self-image, problem-solving abilities, and career orientation. Participants in short, career-oriented workshops who later signed up for courses were compared to those who did not. The instruments were developed by the evaluators for the affective measures, and a portion of the University of Washington entrance exam was used to test problem-solving abilities. The only significant positive change was in self-esteem. Although the authors cite some methodological flaws in their study, they suggest that the program should reconsider both its goals and methods.

Title: "Women's Studies as Change Agent"
Authors: Ruth Scott, Ann Richards, and Marie Wade
Source: *Psychology of Women Quarterly*, Vol. 1, No. 4, Summer 1977

The authors report a study of changes in attitudes toward women by those enrolled in women's studies courses. Participants in the study were enrolled in women's studies courses at a liberal arts college and at a regional campus of a state university. A comparison group consisted of psychology

students at the liberal arts college. The research instrument was the Spence and Helmreich Attitudes Toward Women Scale. Results showed that attitudes toward women became more liberal after taking a women's studies course while there was no change in the comparison group's attitude.

Title: "Women's Studies at Utah State University: A Proposal"
Author: Judith M. Gappa
Source: Educational Resources Information Center, ED 145 822*

The author reports in detail the process of developing a proposal for a women's studies program. This process began with an on-campus needs assessment and moved to a proposal for a full women's studies program. The proposal included recommended evaluation procedures.

Gappa surveyed faculty and students to determine how they viewed changes in sex roles for the future and the effectiveness of the curriculum for preparing students for those roles. She also developed a complete list of women's studies courses that were available or being developed on campus; a list of courses that could be revised to include a women's studies component; and independent studies, research, and fieldwork opportunities relating to women.

The next step in her process was to develop goals and objectives for a total program—and the relevant evaluative approach. The remainder of this summary lists some of those goals and objectives with examples of Gappa's evaluation indicators.

GOAL I:
Develop a sufficient number of courses so students have the options of *electives* or an *area studies concentration of emphasis* within liberal arts.
Objectives
A. Coordinate existing offerings to avoid duplication.
Anticipated outcomes:
1. Total number of courses offered each academic year will be evenly distributed across quarters and will not meet at the same time;
2. Enrollment will increase.
Procedure: Use a form that can be completed from the catalog and includes a space to note unanticipated outcomes.
B. Develop and implement procedures for accepting new courses.
Anticipated outcomes:
1. Proposed courses will follow the procedures of the department, college, and educational policy committee;
2. The women's studies coordinator will advise and assist.
Procedure: Use a worksheet that includes columns for content, instructional style, student evaluation, and faculty comments.

C. Provide a variety of learning experiences for students.
Anticipated outcome: 12-15 departments will offer 50-60 credits in Women's Studies.
Procedure: Use the same worksheet as in B.

GOAL II:
Support efforts in all departments to improve curriculum with respect to women so that objective, nonstereotyped material is presented.
Objectives
A. Increase research.
B. Disseminate material.
C. Undergo careful department evaluation of the quality of research and curriculum offerings.

GOAL III:
Assist the Women's Center.

Goal IV:
Encourage participation of men.

(The last three goals included the same categories of anticipated outcomes and evaluation procedures as the first two.)

The article includes an excellent annotated bibliography.

APPENDIX C: RESOURCE MATERIALS

References on Women's Studies

*Documents with an ED number are contained in the Educational Resources Information Center (ERIC) collection. Microfiche or paper copies of these documents can be obtained for a fee from the ERIC Document Reproduction Service (EDRS), P.O. Box 190, Arlington, Virginia 22210.

American Institutes for Research. *Measures of Educational Equity for Women: A Research Monograph.* American Institutes for Research, 1979. (Available from American Institutes for Research, P.O. Box 1113, Palo Alto, California 94302).

American Institutes for Research. *Sourcebook on Measures of Women's Educational Equity.* 1979. (Available from the Education Development Center, Inc., 55 Chapel Street, Newton, Massachusetts, 02160.)

American Institutes for Research. *WEEA Evaluation Handbook.* 1979.

Bose, Christine, John Steiger, and Philomina Victorine. "Evaluation: Perspectives of Students and Graduates," *Women's Studies Newsletter,* Fall 1977, Vol. V, No. 4, pp. 6-7.

Bose, Christine E., and Janet Priest-Jones. *The Relationship Between Women's Studies, Career Development, and Vocational Choice.* Washington, D.C.: National Institute of Education, in press. (Available from the Social Process/Women's Research Program, National Institute of Education, Washington, D.C. 20208.)

Brush, Lorelei, Alice Ross Gold, and Marni Goldstein White. "The Paradox of Intention and Effect: A Women's Studies Course," *Signs,* 1978, Vol. 3, No. 4.

Canty, Eileen M. "Effects of Women's Studies Courses on Women's Attitudes and Goals," ED 150 490,* 1977.

Davis, Patricia, et al. "GCR 66: Images of Women in Literature, University of Minnesota," ED 124 240,* 1976.

De Biasi, Glen L., and Carolyn Rhodes. "Attitudes of Students, Faculty and Administrators Toward a University Women's Study Program." Old Dominion University, Norfolk, Virginia 23508. (Mimeographed.)

Elovson, Alana Cummings, and Irene Cockroft. "The Reported Impact of Women's Studies Courses on Students' Lives," ED 160 944,* 1977.

Elovson, Alana Cummings. *Women's Studies in Community Colleges.* Washington, D.C.: National Institute of Education, in press. (Available from the Social Process/Women's Research Program, National Institute of Education, Washington, D.C. 20208.)

Gappa, Judith M. "Women's Studies at Utah State University: A Proposal," ED 145 822.*

Green, Mareyjoyce. *The Involvement of Minority Women in Women's Studies.* Washington, D.C.: National Institute of Education, in press. (Available from Social Process/Women's Research Program, National Institute of Education, Washington, D.C. 20208).

Guttentag, Marcia, Lorelei R. Brush, Alice Ross Gold, Marnie W. Mueller, Sheila Tobias, and Marni Goldstein White. "Evaluating Women's Studies: A Decision-Theoretic Approach," *Signs,* 1978, Vol. 3, No. 4.

Hersh, Blanche Glassman. *Re-Entry Women Involved in Women's Studies Courses and Programs.* Washington, D.C.: National Institute of Education, in press. (Available from the Social Process/Women's Research Program, National Institute of Education, Washington, D.C. 20208.)

Howe, Florence. *Seven Years Later: Women's Studies Programs in 1976.* The National Advisory Council on Women's Educational Programs, Washington, D.C.: June 1977.

Howe, Florence, and Paul Lauter. *The Impact on the Institution of Women's Studies Courses and Programs.* Washington, D.C.: National Institute of Education, in press. (Available from the Social Process/

Women's Research Program, National Institute of Education, Washington, D.C. 20208).

Kritek, Phyllis, and Laurie Glass. "Nursing: A Feminist Perspective," *Nursing Outlook*, March 1978.

Mathews, Walter M., Lisa Hunter, and Margaret Robinson. *Developing Successful Proposals in Women's Educational Equity*, Vol. I, II, and III. San Francisco: Far West Laboratory for Educational Research and Development, 1979.

Nelson, Elizabeth Ness, and Kathryn H. Brooks. *Women's Studies as a Catalyst for Faculty Development*. Washington, D.C.: National Institute of Education, in press. (Available from the Social Process/Women's Research Program, National Institute of Education, Washington, D.C. 20208.)

Porter, Nancy. "Evaluation: Reflections of a Program Consultant," *Women's Studies Newsletter*, Fall 1977, Vol. V, No. 4, pp. 3-6.

Porter, Nancy M., and Margaret T. Eileenchild. *The Effectiveness of Women's Studies Teaching*. Washington, D.C.: National Institute of Education, in press. (Available from the Social Process/Women's Research Program, National Institute of Education, Washington, D.C. 20208.)

Reuben, Elaine, and Mary Jo Strauss. *Women's Studies Graduates*. Washington, D.C.: National Institute of Education, in press. (Available from the Social Process/Women's Research Program, National Institute of Education, Washington, D.C. 20208.)

Rosenwood, Linda M., and Patricia Lunneborg. "Women in Continuing Education," ED 067 997,* 1972.

Sackmary, Benjamin, and Hannah Hedrick. "Assessment of the Experiential Learning for College Credit in the Area of Women's Studies," ED 155 208.*

Scott, Ruth, Ann Richards, and Marie Wade. "Women's Studies as Change Agent," *Psychology of Women Quarterly*, Vol. 1, No. 4, Summer 1977.

Shueman, Sharon A., and William E. Sedlacek, "An Evaluation of a Women's Studies Program," *Journal of NAWDAC*, Fall 1977.

References on Evaluation

Abt, Clark, ed. *The Evaluation of Social Programs.* Beverly Hills: Sage Publications, Inc., 1976.

Alkin, Marvin C. "Evaluation Theory Development," *UCLA CSE Evaluation Comment,* No. 2, 1969, pp. 2-7.

American Institutes for Research. *Evaluative Research: Strategies and Methods.* Pittsburgh: American Institutes for Research, 1970.

Anderson, Scarvia B., Samuel Ball, Richard T. Murphy and Associates. *Encyclopedia of Educational Evaluation, Concepts of Technology for Evaluating Educational and Training Programs.* San Francisco: Jossey-Bass, 1973 and 1975.

Apple, Michael W., Michael J. Subkoviak, and Henry S. Lufler, Jr., eds. *Education Evaluation: Analysis and Responsibility.* Berkeley: McCutchan Publishing Co., 1974.

Backstrom, Charles, and Gerald Hursh-Cesar, *Survey Research,* Second Edition. New York: John Wiley and Sons, in press.

Bennett, C. A., and A. A. Lumsdaine, eds. *Evaluation and Experiment: Some Critical Issues in Assessing Social Programs.* New York: Academic Press, 1975.

Bloom, B. S., J. T. Hastings, and G. F. Madaus. *Handbook of Formative and Summative Evaluation.* New York: McGraw-Hill, 1971.

Briloff, Abraham. *Unaccountable Accounting.* New York: Harper and Row, 1973.

California State Department of Education. *California Evaluation Improvement Project,* 1975. (Workshop Materials.)

Campbell, Donald T., and Julian C. Stanley. *Experimental and Quasi-Experimental Designs for Research.* Chicago: Rand McNally, 1963.

Caro, F. G., ed. *Readings in Evaluation Research.* New York: Russell Sage, 1971.

Cook, Thomas, "The Potential and Limitations of Secondary Evaluation," *Education Evaluation: Analysis and Responsibility.* Michael W. Apple, Michael J. Subkoviak, and Henry S. Lufler, eds. Berkeley: McCutchan Publishing Co., 1974.

Cronbach, L. J. "Evaluation for Course Improvement," *New Curricula*. R. W. Heath, ed. New York: Harper, 1963.

Dressel, P. L., et al. *Evaluation in Higher Education*. Boston: Houghton Mifflin, 1961.

Dressel, P. L., and L. B. Mayhew. *General Education—Explorations in Evaluation*, Washington, D.C.: American Council on Education, 1954.

Eash, Maurice J. "Assessing Curriculum Materials: A Preliminary Instrument," *Educational Product Report*, February 1969, Vol. 2, No. 5.

"Evaluation/USOE Manual," *Nation's Schools*, May 1966, Vol. 77, No. 5.

Evaluation, a Journal for Human Service Decision Makers. (Available from *Evaluation*, 501 South Park Avenue, Minneapolis, Minnesota 55415.)

Filstead, W. *Qualitative Methodology*. Chicago: Markham, 1970.

Gardner, Don. "Five Evaluation Frameworks," *The Journal of Higher Education*, September/October 1977, Vol. XLVIII, No. 5, pp. 511-593.

Gronlund, Norman E. *Measurement and Evaluation in Teaching*, second edition. New York: The MacMillan Co., 1971.

Guttentag, Marcia, and Elmer L. Struening, eds. *Handbook of Evaluation Research*, Vols. I and II. Beverly Hills: Sage Publications, Inc., 1975.

Guttentag, Marcia, ed. *Evaluation Studies Review Annual*. Beverly Hills: Sage Publications, Inc., 1977.

Hawkridge, D., P. Campeau, and P. Trickett. *Preparing Evaluation Reports of Educational Programs: A Guide for Authors*. Pittsburgh: American Institutes for Research, 1969.

Heilman, J. G. *Evaluation: A Practical Guide for Evaluators of Social Action Projects*. The Office of Public Service and Research, Auburn (Alabama) University, 1977.

Hemphill, John K. "The Relationships Between Research and Evalua-

tion Studies," *Educational Evaluation: New Roles, New Means*. Ralph Tyler, ed. Chicago: University of Chicago Press, 1969.

House, Ernest R., ed. *School Evaluation: The Politics and Process*. Berkeley: McCutchan Publishing Co., 1973.

Interviewer's Manual, revised edition. Survey Research Center, Institute for Social Research, The University of Michigan, Ann Arbor, Michigan 48106, 1976.

Isaac, S. *Handbook in Research and Evaluation*. San Diego: Robert R. Knapp, Publisher, 1971.

Joint Committee on Evaluation Standards. *Standards for Education Evaluation*. Prepared for the Lily Foundation, the National Science Foundation, and the National Institute of Education, 1978. (Mimeographed.) p. 101.

Klein, Stephen P. *Participant's Notebook: Evaluation Workshop I*. Monterey: CTB/McGraw-Hill Book Company.

Kratwohl, David R. *How to Prepare a Research Proposal: Suggestions for Seeking Funds for Behavioral Science Research*. Syracuse: 1976/77.

McCall, G., and J. Simmons. *Issues in Participant Observation*. Reading, Massachusetts: Addison-Wesley, 1969.

Morris, Lynn Lyons, Carol Taylor Fitz-Gibbon, and Marlene E. Henerson, *Program Evaluation Kit* (eight volumes). Beverly Hills: Sage Publications, Inc., 1978.

National Study of Secondary School Evaluation. *Evaluative Criteria*, 1969 edition. Washington, D.C.: National Study of Secondary School Evaluation, 1969.

Nelson, H., S. Lundin, and F. Biannotta. *Multidisciplinary Evaluation of Educational Innovations: An Anthropological Perspective*. (Paper presented at the annual meeting of the American Educational Research Association, Chicago, 1974.)

Parlett, Malcolm, and Garry Rearden, eds. *Introduction to Illuminative Evaluation*, Pacific Soundings Press, 1977

Popham, James W. *Educational Evaluation*. Englewood Cliffs, New Jersey: Prentice Hall, Inc., 1975.

Popham, James W., ed. *Evaluation in Education: Current Application* Berkeley: McCutchan Publishing Co., 1974.

A Practical Guide to Measuring Project Impact on Student Achievements (Number 1 in a series of monographs on evaluation in education). Washington, D.C.: U.S. Government Printing Office, 1975 (USGPO Stock No. 017-080-01460-2; $1.90).

Provus, Malcolm. "Evaluation of Ongoing Programs in the Public School System," *The 68th Yearbook of the National Society for the Study of Education*, Part II. Chicago: The University of Chicago Press, 1969. pp. 242-283.

Rheinhard, Diane L. *Methodology Development for Input Evaluation Using Advocate and Design Teams.* (Unpublished doctoral dissertation, Ohio State University, 1972.)

Rossi, P. H., and W. Williams, eds. *Evaluating Social Programs.* New York: Seminar Press, 1972.

Scriven, Michael. *Evaluation Bias and Its Control.* Berkeley: University of California, June 1975.

Scriven, Michael. "Evaluation Perspectives and Procedures," *Evaluation in Education: Current Application.* James W. Popham, ed. Berkeley: McCutchan Publishing Co., 1974.

Scriven, Michael. "The Methodology of Evaluation," *Perspectives of Curriculum Evaluation.* Ralph W. Tyler, ed. Chicago: Rand McNally, Inc., 1967.

Shoemaker, D. M. "Evaluating the Effectiveness of Competing Instructional Programs," *Educational Researcher*, May 1972, Vol. I.

Sieber, Samuel D. *Formative Evaluation: An Exploration with Case Materials.* New York: Bureau of Applied Social Research, Columbia University, 1972.

Smith, L., ed. *Anthropological Perspectives on Evaluation.* AERA Monograph Series in Evaluation. Chicago: Rand McNally, 1974.

Stake, Robert, and Craig Gjerde. *An Evaluation of TCITY: The Twin City Institute for Talented Youth.* The Center for Instructional Research and Curriculum Evaluation, University of Illinois, 1971.

Stake, Robert E. "The Countenance of Educational Evaluation," *Teachers College Record*, 1967, Vol. 68.

Steele, Sara. *Contemporary Approaches to Program Evaluation*. ERIC Clearinghouse on Adult Education. Washington, D.C.: Capitol Publications, 1973.

Stufflebeam, Daniel, et al. *Educational Evaluation and Decision Making*. Itasca, Illinois: F. E. Peacock Publishers, 1971.

Stufflebeam, Daniel L. "Meta Evaluation," *The Evaluation Center Occasional Paper Series*, 3, Western Michigan University, May 1975.

Stufflebeam, Daniel L. "Evaluation as Enlightenment for Decision-Making," Ohio State University Evaluation Center, 1968. (Mimeographed.)

Suchman, E. A. *Evaluation Research: Principles and Practice in Public Service and Social Action Programs*. New York: Russell Sage, 1967.

Talburtt, Margaret A. *Evaluation of Project WELD: Women's Education Learning and Doing*. Ann Arbor, Michigan: Formative Evaluation Research Associates, 1978.

Tyler, Ralph W. "General Statement on Evaluation," *Journal of Educational Research*, Vol. 35, 1942.

Tyler, Ralph W. et al., eds., *Perspectives in Curriculum Evaluation*. Chicago: Rand McNally, 1967.

Walberg, Herbert J., ed., *Evaluating Educational Performance: A Sourcebook on Methods, Instruments, and Examples*. Berkeley: McCutchan Publishing Co., 1974.

Webb, Eugene, Donald Campbell, and Richard D. Schwartz. *Unobtrusive Measures: Nonreactive Research in the Social Sciences*. Chicago: Rand McNally, 1966.

Williams, Martha K., et al. *Educational Manager's Guide to Project Evaluation*. Andover, Massachusetts: The Network, 1978.

Wilson, Stephen. "The Use of Ethnographic Techniques in Educational Research," *Review of Educational Research*, Spring 1977, Vol. 47, No. 2.

Wolf, Robert L. *The Application of Select Legal Concepts to Education Evaluation.* (Unpublished doctoral dissertation, University of Illinois, 1973.)

Worthen, Blaine R. *Toward a Taxonomy of Evaluation Designs.* Columbus: The Evaluation Center, Ohio State University, 1968. (Paper read at the 1968 Annual Meeting of the American Educational Research Association, Chicago, February 1968.)

References on Technical Issues

American Psychological Association. *Standards for Educational and Psychological Tests.* Washington, D.C.: American Psychological Association, 1974.

Anderson, R. C. "How to Construct Achievement Tests to Assess Comprehension," *Review of Educational Research,* 1972, Vol. 42.

Ebel, R. L. *Measuring Educational Achievement,* (second edition). Englewood Cliffs, New Jersey: Prentice Hall, 1972.

Ferguson, G. A. *Statistical Analysis in Psychology and Education.* New York: McGraw-Hill, 1966.

Popham, James W. *Educational Statistics: Use and Interpretation.* New York: Harper and Row, 1967.

Simon, A., and E. G. Boyer, eds. *Mirrors for Behavior: An Anthology of Classroom Observation Instruments.* Philadelphia: Research for Better Schools, Inc., 1967 and 1971 (two volumes).

Tyler, Ralph W. *Constructing Achievement Tests.* Columbus, Ohio: Ohio State University, 1934.

• U. S. GOVERNMENT PRINTING OFFICE : 1980 630-972/827